Properties of damped Ly α absorption systems and star-forming galaxies in semi-analytic models at z = 2

Berry, Michael; Somerville, Rachel S.; Gawiser, Eric; Maller, Ariyeh H.; Popping, Gergö; Trager, Scott C.

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
10.1093/mnras/stw231

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **458,** 531–557 (2016)
Advance Access publication 2016 February 17

doi:10.1093/mnras/stw231

Properties of damped Ly α absorption systems and star-forming galaxies in semi-analytic models at $z = 2$

Michael Berry,[1] Rachel S. Somerville,[1]★ Eric Gawiser,[1] Ariyeh H. Maller,[2,3] Gergö Popping[4,5] and Scott C. Trager[5]

[1]*Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854-8019, USA*
[2]*Physics Department, New York City College of Technology, Brooklyn, NY 11201, USA*
[3]*Department of Astrophysics, American Museum of Natural History, New York, NY 10024, USA*
[4]*European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748, Garching, Germany*
[5]*Kapteyn Astronomical Institute, University of Groningen, Postbus 800, NL-9700 AV Groningen, the Netherlands*

Accepted 2016 January 26. Received 2016 January 25; in original form 2015 April 27

ABSTRACT

We investigate predictions from semi-analytic cosmological models of galaxy formation for the properties of star-forming galaxies (SFGs) and damped Ly α absorption systems (DLAS), and the relationship between these two populations. Our models reproduce fairly well the observed distributions of redshift, stellar mass, star formation rate (SFR), and dust extinction for $z \sim 2$ SFGs. We predict that DLA hosts span a broad range of properties, with broad and relatively flat distributions of stellar and halo mass, SFR, and luminosity. The photometric colours of DLA host galaxies trace the colours of galaxies with similar luminosities, but the majority are much fainter than the limits of most existing surveys of SFGs. Generally, DLA host galaxies and SFGs at $z = 2$ follow similar trends between stellar mass, DLA cross-section, cold gas fraction, SFR, metallicity, and dust extinction as the global population of galaxies with the same stellar mass. Since DLAS select galaxies with larger cold gas masses, they tend to have larger cold gas fractions, lower metallicities, higher SFRs, and less dust extinction than galaxies at the same stellar mass. Our models reproduce the observed relations between impact parameter, column density, and metallicity, suggesting that the sizes of the gas discs giving rise to DLAS in our models are roughly correct. We find that molecular fractions and SFRs are in general significantly lower at the location of the DLA line of sight than the galaxy-averaged value.

Key words: galaxies: evolution – galaxies: formation – galaxies: high-redshift – quasars: absorption lines.

1 INTRODUCTION

Damped Ly α systems (DLAS), H I absorption systems with column densities of $N_{\rm H\,I} > 2 \times 10^{20}$ cm^{-2}, contain a substantial fraction of the neutral hydrogen in the Universe (Storrie-Lombardi & Wolfe 2000; Wolfe & Prochaska 2000; Péroux et al. 2003; Noterdaeme et al. 2009, 2012; Prochaska & Wolfe 2009). DLAS at $z \sim 2$ are therefore thought to represent a significant reservoir of gas that can fuel star formation (SF; Storrie-Lombardi & Wolfe 2000). At $1 < z < 3$, coinciding with the peak epoch of star formation, star-forming galaxies (SFGs) boast star formation rates of SFR $\sim$10 to $>$100 M$_\odot$ yr^{-1}, presumably indicating high gas fractions (Shapley et al. 2001, 2005; Reddy et al. 2006; Tacconi et al. 2010, 2013). In the classical picture, DLAS arise from neutral gas in the interstellar medium (ISM) of high-redshift galaxies, and therefore one might expect a strong link between DLAS and SFGs. However, a clear observational link between these two populations has not been found. Moreover, as DLAS are detected in absorption, they provide us with one of the best ways in which to study atomic hydrogen at high redshift. Compared to traditional galaxy surveys, where galaxies are identified based on radiation emitted by stars, selecting galaxies in absorption represents an orthogonal means of studying galaxy formation, as galaxy selection will not be directly dependent on the SFR or stellar mass of the host galaxy. Therefore, understanding the connection between DLAS, SFGs, and the underlying galaxy population will yield insight into the physics of galaxy formation, and should be a key target for galaxy formation models.

Deep and wide surveys have generated large catalogues of SFGs across the redshift range $1.5 \lesssim z \lesssim 3.5$ through a variety of colour

★ E-mail: somerville@physics.rutgers.edu

selection techniques (e.g. Steidel et al. 1996, 2003, 2004; Daddi et al. 2004; Hathi et al. 2010). Extensive follow-up observations and detailed analyses have generated a wealth of information regarding the physical properties of SFGs at high redshift. The most common selection method (Steidel et al. 1996) identifies SFGs based on the location of the Lyman break (called BM/BX selection). This technique has also been applied using space-based photometry to select UV-bright SFGs at $1.4 \lesssim z \lesssim 3$ and at fainter magnitudes (e.g. $B_{435} < 26.5$ for U275-band dropouts at $z \sim 2$; Hathi et al. 2010; Oesch et al. 2010). The 'BzK' selection technique (Daddi et al. 2004; Reddy et al. 2005) has also been used to identify galaxies at $1.4 \lesssim z \lesssim 2.5$ selected in the K band. This technique is thought likely to be closer to a 'stellar mass' selection, and less likely to miss objects with red colours, whether due to old stellar populations or dust.

It is known that there is a large overlap between the populations selected using these different methods, e.g. ∼70 per cent of BzKs are also BM/BX galaxies (Grazian et al. 2007). The advent of medium-band surveys and accurate photometric redshifts has allowed the identification of a large number of galaxies across a wide range of redshifts (e.g. Whitaker et al. 2011; Dahlen et al. 2013), generating flux-limited samples of galaxies. Studies of high-redshift galaxy properties are being increasingly conducted with magnitude-limited photo-z samples rather than colour-selected samples, however given the very large literature based on colour-selected samples, it is interesting to ask how these different selection techniques trace the underlying galaxy population in mock samples drawn from theoretical models.

The identification of large numbers of optically-detected quasars (QSOs) at $z > 2$ in the Sloan Digital Sky Survey has produced a substantial catalogue of absorbers at high redshift. In spite of extensive efforts, finding galaxy counterparts to DLAS has been challenging, yielding only a handful of host galaxies (Warren et al. 2001; Krogager et al. 2012). As a result, establishing a connection between DLAS and emission-selected galaxies has remained an open challenge. Observations of [C II] 158 μm cooling rates in DLAS exhibit a bimodality, with some DLAS consistent with trace amounts of SF and others likely associated with substantial SF (Wolfe et al. 2008, hereafter W08). In fact, W08 find that the low cooling rate subset is also consistent with being heated by the extragalactic background light. These DLAS with low cooling rates may arise in low-mass haloes with inefficient SF, or intergalactic or circumgalactic gas. On the other hand, the DLAS with high cooling rates are typically found in more massive dark matter haloes as tracked by their velocity widths, Δv_{90}, which are assumed to correlate with virial velocity. These DLAS also exhibit higher metallicities, higher dust-to-gas ratios, and higher star formation rate (SFR) surface densities ($\Sigma_{\rm SFR}$). Curiously, their H I column density distributions are similar to the low cooling rate subset. Due to the high cooling rates and infrequent occurrence of extended low surface brightness galaxies (e.g. Wolfe & Chen 2006), W08 comment that the high cooling rate DLAS are likely heated by far-ultraviolet radiation from a nearby bulge source, even though few DLA host galaxies have been observed. An updated study of cooling rates in DLAS has been presented in Neeleman et al. (2013).

A natural explanation for these DLAS is that they arise from cold neutral gas surrounding a star-forming region, such as a nearby SFG that is beyond our current detection limits. Further constraints on the SF efficiency of the neutral gas contained in DLAS has been obtained by statistical comparisons between the number of SFGs predicted from the observed DLA population under a given set of assumptions and the actual detected SFG population. Wolfe & Chen (2006) derived upper limits on the SF efficiency in DLAS using the Hubble Ultra-deep Field (HUDF). They found that the normalization of the Kennicutt relation (Kennicutt 1998) in DLAS must be at least a factor of 30–100 lower at $z \sim 3$ in regions where no SFGs were detected at the depth of the HUDF V-band image. Rafelski, Wolfe & Chen (2011) performed a similar comparison, again using the HUDF, but using the outer regions of detected Lyman Break Galaxies (LBG). They similarly inferred very low SF efficiencies in DLAS, 10–50 times lower than the expectation from the local Kennicutt relation. The recent study of Fumagalli et al. (2015) used direct imaging of 'double DLA' systems (in which the glare of the quasar is eliminated by a second DLAS along the line of sight) and similarly found low SFR densities in DLAS, consistent with the statistical constraints described above and suggesting that DLA are not associated with bright SFGs.

Cooke et al. (2006) found an overdensity of LBGs and QSOs around DLAS at $z \sim 3$, indicating a possible connection between the size and density of systems with high H I column densities and sources of significant ionizing radiation. Using numerical simulations, Bouché et al. (2005) measured the clustering of DLAS and found them to occupy moderate mass haloes with an upper limit of $\log M_{\rm h}/{\rm M}_\odot \sim 11.1$. Moreover through a cross-correlation analysis of DLAS and the Lyman α forest, Font-Ribera et al. (2012) find that DLAS populate relatively massive haloes of $M_{\rm h} = 6 \times 10^{11}\,{\rm M}_\odot$. These results indicate that some fraction of DLAS are likely associated with SFGs, yet the absence of detected DLA galaxy counterparts and the upper limits on associated SF in DLAS may suggest that a substantial fraction arise from a separate populations of objects.

In spite of this tension, there has been relatively little theoretical work to date aimed at studying the relationship between DLAS and SFGs. In their numerical hydrodynamic 'zoom' simulations, Pontzen et al. (2008) found that DLAS at $z = 3$ typically reside in lower mass haloes ($10^9 < M_{\rm h}/{\rm M}_\odot < 10^{11}$) with very low SFRs (${\rm SFR} < 0.1\,{\rm M}_\odot\,{\rm yr}^{-1}$) and stellar masses ($10^{6.5} < M_*/{\rm M}_\odot < 10^{9.5}$). In contrast, using a larger volume cosmological hydrodynamic simulation, Cen (2012) found DLAS to reside in more massive galaxies ($10^{10} < M_{\rm h}/{\rm M}_\odot < 10^{12}$ at $1.6 < z < 4$), and found that galactic winds play a crucial role in DLA kinematics. They find DLAS hosts at $3 < z < 4$ to have SFRs ∼0.3−30 ${\rm M}_\odot\,{\rm yr}^{-1}$ and a 20–30 per cent overlap fraction with LBGs. The majority were found to be further than 100 kpc away from L^* galaxies, suggesting that they are associated with gas that is not experiencing rapid SF. Recently, Rahmati & Schaye (2014) found an anticorrelation between H I column density and impact parameter where impact parameter increases with the mass of the host galaxy. They also frequently found the galaxies nearest to the DLA to be too faint to be detected in current surveys, suggesting that DLA hosts may frequently be misidentified, and leading to a systematic overestimate in impact parameter and host mass.

A concern since the early days of DLA studies is the possible bias that could arise if the intervening absorber causes the QSO to become too faint or too reddened to be included in observed samples. Numerous studies have sought to characterize the amount of dust obscuration and its potential bias on the selection of DLAS (Pei, Fall & Bechtold 1991; Ellison et al. 2001; Vladilo, Prochaska & Wolfe 2008; Khare et al. 2012). One common technique is to compare the stacked colours of quasars with DLAS along the sightline to a control sample. Historically, this technique has indicated very small amounts of dust reddening, $E(B-V) \lesssim 0.05$ (e.g. Murphy & Liske 2004; Ellison, Hall & Lira 2005). Recently, Khare et al. (2012) investigated a sample of 1084 absorption systems

with log $(N_{\rm H\,I}) > 20$ cm^{-2} at $2.15 < z < 5.2$ and found that $\lesssim$ 10 per cent have relatively high chemical abundances and cause significant reddening, leading them to conclude that about 10 per cent of DLAS could be missed in colour-selected QSO surveys. Furthermore, Ellison et al. (2001) compiled a sample of radio-selected QSOs irrespective of optical magnitude and found the number density per unit redshift of DLAS to be higher towards fainter targets. They concluded that this dust-induced bias indicates that the cosmological H I mass density is likely underestimated along with the number density of DLAS, but by no more than a factor of 2 at $z \sim 2.3$. Finally, Jorgenson et al. (2006) report the existence of fields with radio-selected QSOs with no optical counterparts, indicating the presence of highly obscured QSOs, possibly due to dusty foreground galaxies. These studies suggest that the majority of DLAS have negligible amounts of dust extinction while some unknown fraction may contain significant amounts of dust, leading to the underestimation of the amount of H I in the Universe.

One approach to understanding the connection between DLAS and galaxies detected in emission is through the use of semi-analytic models (SAMs). SAMs are based on the cold dark matter (CDM) paradigm for structure formation (Blumenthal et al. 1984), and have been widely used to produce a general picture of how density fluctuations in the primordial universe evolve into the observable galaxy population. They use simple but physically motivated 'recipes' to track bulk quantities such as the total mass in stars, hot gas, cold gas, metals within a galaxy, the dark matter halo, the halo infall region, and the intergalactic medium. Numerical hydro simulations offer the advantage of detailed spatial and kinematic information, but also rely heavily on 'sub-grid' recipes for important processes such as SF and stellar feedback. On the other hand, SAMs allow one to vary the details of these recipes to explore the associated uncertainties and allow the production of 'mock catalogues' of very large numbers of galaxies, which is still inaccessible for numerical hydro simulations.

We use the 'Santa Cruz' SAM first presented in Somerville & Primack (1999) and recently updated to include separate tracking of atomic, molecular, and ionized gas, and an H_2-based SF recipe (Popping, Somerville & Trager 2014; Somerville, Popping & Trager 2015, hereafter PST14 and SPT15, respectively). These models do not explicitly resolve sub-galactic structure, but assume that gas discs have exponential radial profiles with scalelengths estimated based on the angular momentum of the gas. This model has been shown to reproduce the evolution of disc sizes from $z \sim 2$ to 0 (Somerville et al. 2008b), the observed sizes of H I discs in the local universe, the observed sizes of CO discs in local and high-redshift galaxies, and the spatial extent of the SFR density in nearby and high-redshift galaxies (PST14).

In Berry et al. (2014, hereafter B14), we presented the predictions of these models for the properties of DLAS at $0 < z < 5$, and explored the effect of different gas configurations and H_2 formation recipes. We investigated two H_2 formation recipes: one recipe is based on the empirical relationship between H_2 fraction and disc mid-plane pressure (Blitz & Rosolowsky 2004, 2006, hereafter BR), and the other is based on fitting functions from hydrodynamic simulations, in which H_2 fraction depends on the cold gas surface density, cold gas metallicity, and the local UV background (Gnedin & Kravtsov 2011, hereafter GK). Both H_2 formation recipes produced similar predictions for DLAS as well as similar stellar mass functions, SFR distributions, etc. although the GK model performed slightly better for DLA properties. Regardless of the gas partitioning recipe, we found that a 'standard disc' model, in which the average specific angular momentum of the gas that forms the disc is the same as that of the dark matter halo, significantly underproduced the number of DLAS and sub-DLAS at column densities $\log N_{\rm H\,I} \lesssim 21.2$. The 'standard disc' models also produced poor agreement with observed DLA kinematics and with constraints on halo masses from clustering (Font-Ribera et al. 2012). As a result, we explored an 'extended gas' model, in which we assumed that the average specific angular momentum of the gas is 2.5 times larger than that of the dark matter halo. These extended gas models produced fairly good agreement with the observed H I column density distribution down to $\log N_{\rm H\,I} \lesssim 19$, and with observed DLA properties at $z \lesssim 2.5$, including metallicities, kinematics, and halo masses. Moreover, the extended gas picture is supported by recent numerical hydrodynamic simulations, which find that stream-fed 'cold accretion' leads to similar extended gas discs with a factor of approximately 2 to 3 larger specific angular momentum than the average value for the dark matter halo (Brooks et al. 2009; Stewart et al. 2013; Danovich et al. 2015).

In this paper, we use these new models to obtain predictions for the properties of SFGs at $z \sim 2$, and to explore the effects of different colour selection criteria in identifying these objects. In addition, we explore the relationship between DLAS and SFGs, their relative properties, and the differences between properties of DLAS as measured along the line of sight and the average for their host galaxies. We restrict our analysis to the redshift range $1.7 < z < 2.3$, where our models are the most successful at reproducing the properties of DLAS. We create lightcones based on the Bolshoi *N*-body simulation (Klypin, Trujillo-Gomez & Primack 2011), which provide the spatial locations of the dark matter haloes that are populated with gas and stars using our SAM. We then identify DLAS by passing lines of sight through these lightcones. We characterize the statistical properties of DLA host galaxies in terms of their luminosities, colours, masses (stellar, cold gas, and total), SFRs, DLA cross-sections, sizes, metallicities, and amount of dust. Through comparing the properties along the DLA line of sight to the average DLA host galaxy properties, we hope to obtain potential insights into the interpretation of DLA observations and their connection with SFGs. Finally, we use our catalogues of DLAS and SFGs to study the relations between stellar mass and DLA cross-section, cold gas fraction, SFR, metallicity, and dust reddening as traced by galaxy scaling relations. Our paper is organized as follows. In Section 2, we briefly describe the main ingredients in the models. Sections 3 and 4 describe how we select DLAS and SFGs, and present a comparison between model and observed SFGs. We present the predicted properties of DLA host galaxies, compare DLA properties to those of the global galaxy population, and compare DLA host galaxies to SFGs in Section 5. In Section 6, we discuss the implications of our results and the connection with the current literature. Finally, we summarize our conclusions in Section 7.

2 MODELS AND METHODOLOGY

2.1 The SAM of galaxy formation model

The SAMs used here to compute the formation and evolution of galaxies within a ΛCDM cosmology were originally presented in Somerville & Primack (1999) and Somerville, Primack & Faber (2001), with significant updates described in detail in Somerville et al. (2008a, hereafter S08), Somerville et al. (2012, hereafter S12), Porter et al. (2014), and PST14. The Santa Cruz SAM includes the following main ingredients: (1) the growth of dark matter structure in a hierarchical clustering framework as described by 'merger trees', (2) shock heating and radiative cooling of gas, (3) conversion

of cold gas into stars via an empirical 'Kennicutt–Schmidt' (KS) relation, (4) evolution of stellar populations, (5) ejection of cold gas from galaxies by stellar-driven winds (6) metal enrichment of the interstellar and intracluster medium (7) 'quasar' and 'radio' mode black hole (BH) growth and feedback from AGN, (8) starbursts and morphological transformation due to galaxy mergers. Here, we briefly summarize these ingredients – for a more detailed description of the models, see S08, S12, and PST14. Throughout this work, we assume a standard ΛCDM universe ($\Omega_{\rm m} = 0.28$, $\Omega_\Lambda = 0.72$, $H_0 = 70.0$, $\sigma_8 = 0.81$, $n_{\rm s} = 0.96$, baryon fraction $f_{\rm b} = 0.1658$), consistent with the *WMAP7* cosmological constraints, and also with the parameters adopted in the Bolshoi simulation (Klypin et al. 2011; Komatsu et al. 2011). In addition, we adopt a Chabrier stellar initial mass function (IMF; Chabrier 2003). All quoted magnitudes are AB, unless otherwise specified.

We obtain a catalogue of 'root' haloes by extracting haloes along lightcones from the Bolshoi simulations (Behroozi, Conroy & Wechsler 2010; Klypin, Trujillo-Gomez & Primack 2011). These lightcones cover a 1 by 1 deg^2 area on the sky over a redshift range $0 < z < 5$ and contain galaxies with dark matter halo masses from $10^{9.5}$ to $10^{14.5}$ $\mathrm{M}_\odot$. The Bolshoi simulation starts to become incomplete at $V_{\rm vir} \simeq 50$ km s^{-1}, $M_{\rm h} \simeq 10^{10}$ $\mathrm{M}_\odot$; see Klypin et al. (2011) for more details.

The merging histories (or merger trees) of dark matter haloes are constructed based on the Extended Press-Schechter (EPS) formalism using the method described in Somerville & Kolatt (1999), with improvements described in S08. These merger trees record the growth of dark matter haloes via merging and accretion, with each 'branch' representing a merger of two or more haloes. We follow each branch back in time to a minimum progenitor mass $M_{\rm res}$, which we refer to as the mass resolution of our simulation. We have adopted $M_{\rm res} = 10^{9.5}$ $\mathrm{M}_\odot$ in all the models presented here. Our SAMs give nearly identical results when run on the EPS merger trees or on merger trees extracted from dissipationless *N*-body simulations (Lu et al. 2014; Porter et al. 2014).

Whenever dark matter haloes merge, the central galaxy of the largest progenitor becomes the new central galaxy, and all others become 'satellites'. Satellite galaxies lose angular momentum due to dynamical friction as they orbit and may eventually merge with the central galaxy. To estimate this merger time-scale, we use a variant of the Chandrasekhar formula from Boylan-Kolchin, Ma & Quataert (2008). Tidal stripping and destruction of satellites are also included as described in S08. We have checked that the resulting mass function and radial distribution of satellites (sub-haloes) agrees with the results of high-resolution *N*-body simulations that explicitly follow sub-structure (Macciò et al. 2010).

Before reionization, each halo contains a mass of hot gas equal to the universal baryon fraction times the virial mass of the halo. After reionization, which we assume to be complete by $z = 11$, the photoionizing background suppresses the collapse of gas into low-mass haloes. We use the results of Gnedin (2000) and Kravtsov, Gnedin & Klypin (2004) to model the fraction of baryons that can collapse as a function of halo mass and redshift after reionization.

When a dark matter halo collapses or experiences a merger that more than doubles the mass of the largest progenitor, the hot gas is assumed to be shock-heated to the virial temperature of the new halo. This gas then cools and collapses based on a simple spherically symmetric model. We assume that cold gas is accreted only by the central galaxy, even though realistically, satellite galaxies should receive some fraction of the new cold gas.

All newly cooling gas collapses to form a rotationally supported disc. The scale radius of the disc is computed based on the initial angular momentum of the gas and the halo profile, assuming angular momentum is conserved and the self-gravity of the collapsing baryons causes contraction in the inner part of the halo (Blumenthal et al. 1986; Flores et al. 1993; Mo, Mao & White 1998). Somerville et al. (2008b) showed that this approach reproduced the observed size versus stellar mass relation for discs from $z \sim 0$ to 2, and PST14 showed that it reproduces the spatial extent of H I in low-redshift galaxies, and H_2 (as traced by CO) in galaxies at $z \sim 0$–2. In the 'standard disc' scenario, the cold gas specific angular momentum is set equal to that of the dark matter, i.e. $f_j = j_{\rm gas}/j_{\rm DM} = 1$. In this paper, as in B14, we also consider an 'extended disc' scenario in which $f_j = 2.5$, i.e. the disc material has a specific angular momentum 2.5 times larger than the average value for dark matter in the halo.

SF occurs in two modes, a 'normal' mode in isolated discs, and a merger-driven 'starburst' mode. SF in the 'normal' mode is modelled as described in Section 2.3 below. The efficiency and time-scale of the merger-driven burst mode is a function of the merger mass ratio and the gas fractions of the progenitors, and is based on the results of hydrodynamic simulations of binary galaxy mergers (Robertson et al. 2006a; Hopkins et al. 2009b).

Some of the energy from supernovae and massive stars is assumed to be deposited in the ISM, resulting in the driving of a large-scale outflow of cold gas from the galaxy. The mass outflow rate is parametrized as a function of the galaxy circular velocity times the SFR, as motivated by the 'energy driven' wind scenario.

Some fraction of this ejected gas escapes from the potential of the dark matter halo, while some is deposited in the hot gas reservoir within the halo, where it becomes eligible to cool again. The fraction of gas that is ejected from the disc but retained in the halo versus ejected from the disc and halo is a function of the halo circular velocity (see S08 for details), such that low-mass haloes lose a larger fraction of their gas. The gas that is ejected from the halo is kept in a larger 'reservoir', along with the gas that has been prevented from falling in due to the photoionizing background. This gas is allowed to 're-accrete' on to the halo as described in S08.

Each generation of stars also produces heavy elements, and chemical enrichment is modelled in a simple manner using the instantaneous recycling approximation. For each parcel of new stars $\mathrm{d}m_*$, we also create a mass of metals $\mathrm{d}M_Z = y\,\mathrm{d}m_*$, which we assume to be instantaneously mixed with the cold gas in the disc. The yield y is assumed to be constant, and is treated as a free parameter. When gas is removed from the disc by supernova-driven winds as described above, a corresponding proportion of metals is also removed and deposited either in the hot gas or outside the halo, following the same proportions as the ejected gas. Metals also 're-accrete' into the halo at the same rate as the ejected gas.

Mergers are assumed to remove angular momentum from the disc stars and to build up a spheroid. The efficiency of disc destruction and spheroid growth is a function of progenitor gas fraction and merger mass ratio, and is parametrized based on hydrodynamic simulations of binary galaxy mergers (Hopkins et al. 2009b). These simulations indicate that more 'major' (closer to equal mass ratio) and more gas-poor mergers are more efficient at removing angular momentum, destroying discs, and building spheroids. Note that the treatment of spheroid formation in mergers used here has been updated relative to S08 as described in Hopkins et al. (2009a). The updated model produces good agreement with the observed fraction of disc versus spheroid-dominated galaxies as a function of stellar mass (Hopkins et al. 2009a; Porter et al. 2014).

In addition, mergers drive gas into galactic nuclei, fuelling BH growth. Every galaxy is born with a small 'seed' BH (typically

$\sim 100\,M_\odot$ in our standard models). Following a merger, any pre-existing BHs are assumed to merge quickly, and the resulting hole grows at its Eddington rate until the energy being deposited into the ISM in the central region of the galaxy is sufficient to significantly offset and eventually halt accretion via a pressure-driven outflow. This results in self-regulated accretion that leaves behind BHs that naturally obey the observed correlation between BH mass and spheroid mass or velocity dispersion (Di Matteo, Springel & Hernquist 2005; Robertson et al. 2006b; S08).

There is a second mode of BH growth, termed 'radio mode', that is thought to be associated with powerful jets observed at radio frequencies. In contrast to the merger-triggered mode of BH growth described above (sometimes called 'bright mode' or 'quasar mode'), in which the BH accretion is fuelled by cold gas in the nucleus, here, hot halo gas is assumed to be accreted according to the Bondi–Hoyle approximation (Bondi 1952). This leads to accretion rates that are typically only about $\lesssim 10^{-3}$ times the Eddington rate, so that most of the BH's mass is acquired during episodes of 'bright mode' accretion. However, the radio jets are assumed to couple very efficiently with the hot halo gas, and to provide a heating term that can partially or completely offset cooling during the 'hot flow' mode (we assume that the jets cannot couple efficiently to the cold, dense gas in the infall-limited or cold flow regime).

2.2 Multiphase gas partitioning

Throughout this paper we refer rather loosely to 'cold' gas, which is gas that according to our simple cooling model has been able to cool below 10^4 K via radiative atomic cooling. Most previous cosmological simulations have considered all of this 'cold' gas to be eligible to form stars. Here, we partition it into components that we label atomic, molecular, and ionized, and only allow the 'molecular' component to participate in SF. As we do not explicitly track the temperature or density of the 'cold' gas in our models, this is obviously still extremely schematic. However, when we refer to 'cold' gas, we are referring to gas that is in one of these three states and is dynamically associated with the galactic disc (rather than in an extended hot halo, an outflow, etc).

At each timestep, we compute the scale radius of the cold gas disc using the angular momentum based approach described above, and assume that the *total* ($H\,I + H_2$) cold gas distribution is described by an exponential with scale radius $r_{\rm gas}$. We do not attempt to track the scale radius of the stellar disc separately, but make the simple assumption that $r_{\rm gas} = \chi_{\rm gas} r_{\rm star}$, with $\chi_{\rm gas} = 1.7$ fixed to match stellar scalelengths at $z = 0$. Bigiel & Blitz (2012) showed that this is a fairly good representation, on average, for the discs of nearby spirals. We then divide the gas disc into radial annuli and compute the fraction of molecular gas, $f_{\rm H_2}(r) \equiv \Sigma_{\rm H_2}(r)/[\Sigma_{\rm H_2}(r) + \Sigma_{\rm H\,I}(r)]$, in each annulus, as described below.

2.2.1 *Ionized gas*

Most previous SAMs have neglected the ionized gas associated with galaxies, which may be ionized either by an external background or by the radiation field from stars within the galaxy. Here, we include a simple analytic estimate of the ionized gas fraction motivated by the model presented in Gnedin (2012). We assume that some fraction of the total cold gas in the galaxy, $f_{\rm ion,int}$, is ionized by the galaxy's own stars. In addition, a slab of gas on each side of the disc is ionized by the external background radiation field. Gas is assumed to be ionized if it lies below a critical threshold surface density $\Sigma_{\rm H\,II}$. Throughout this paper, we assume $f_{\rm ion,int} = 0.2$ (as in the Milky Way) and $\Sigma_{\rm H\,II} = 0.4\,M_\odot\,{\rm pc}^{-2}$, motivated by Gnedin (2012). Applying this model within our SAM gives remarkably good agreement with the ionized fractions as a function of circular velocity shown in fig. 2 of Gnedin (2012), obtained from hydrodynamic simulations with time dependent and spatially variable 3D radiative transfer of ionizing radiation from local sources and the cosmic background.

2.2.2 *Molecular gas*

We consider two approaches for computing the molecular gas fractions in galaxies. The first is based on the empirical pressure-based recipe presented by Blitz & Rosolowsky (2006), who found that the molecular fraction $R_{\rm mol} \equiv \Sigma_{\rm H_2}/\Sigma_{\rm H\,I}$ is correlated with the disc hydrostatic mid-plane pressure P in nearby galaxies. The hydrostatic pressure changes as a function of distance from the centre of the disc based on the vertical velocity dispersions of the gas and stars (Elmegreen 1989, 1993).

The second approach is based on results from high-resolution 'zoom-in' cosmological simulations with the Adaptive Refinement Tree code of Kravtsov, Klypin & Khokhlov (1997), including gravity, hydrodynamics, non-equilibrium chemistry, and simplified radiative transfer (Gnedin & Kravtsov 2010; GK). These simulations are able to follow the formation of molecular hydrogen through primordial channels and on dust grains, as well as dissociation of molecular hydrogen and self- and dust-shielding. We use the fitting function of GK, which parameterizes the fraction of molecular hydrogen as a function of the dust-to-gas ratio relative to the Milky Way, $D_{\rm MW}$, the UV ionizing background relative to the Milky Way $U_{\rm MW}$, and the neutral gas surface density $\Sigma_{\rm H\,I+H_2}$. Following Gnedin & Kravtsov (2010), we take the dust-to-gas ratio to be equal to the metallicity of the cold gas in solar units, $D_{\rm MW} = Z/Z_\odot$. We assume that the local ionizing UV background scales in proportion to the SFR relative to the Milky Way value, $U_{\rm MW} = {\rm SFR}/\,M_\odot\,{\rm yr}^{-1}$. For more details on molecular hydrogen partitioning, see SPT15 and PST14.

The fitting functions from Gnedin & Kravtsov (2010) characterize the formation of H_2 on dust grains, which becomes the dominant mechanism once the gas has reached a few tenths of Solar metallicity. Other channels for H_2 formation in primordial gas must be responsible for producing the molecular hydrogen out of which the first stars formed. Studies with numerical hydrodynamic simulations containing detailed chemical networks and analytic calculations for H_2 formation through primordial channels have shown that 'Pop III' stars pollute their surroundings and enrich the ISM to $Z_{\rm III} \sim 10^{-3}\,Z_\odot$ (e.g. Nakamura & Umemura 2001; Schneider et al. 2002; Greif et al. 2010; Wise et al. 2012; Glover 2013). Since these processes are thought to have taken place in haloes that are much smaller than our resolution limit, we represent them in a simple manner, by setting a 'floor' to the molecular hydrogen fraction in our haloes, $f_{\rm H_2,floor}$. In addition, we 'pre-enrich' the initial hot gas in haloes, and the gas that is accreted on to haloes due to cosmological infall, to a metallicity of $Z_{\rm pre\text{-}enrich}$. We adopt typical values of $f_{\rm H_2,floor} = 10^{-4}$ and $Z_{\rm pre\text{-}enrich} \sim 10^{-3}\,Z_\odot$, based on the numerical simulation results mentioned above (Haiman, Rees & Loeb 1996; Bromm & Larson 2004). As shown in SPT15, the model predictions are not sensitive to reasonable variations of these parameters.

2.3 Star formation

The 'classical' KS recipe (Kennicutt 1998) assumes that the surface density of SF in a galaxy is a function of the *total* surface density of

the cold neutral gas (atomic and molecular), above some threshold surface density $\Sigma_{\rm crit}$. This approach has been used to model SF in most previous SAMs and numerical hydrodynamical simulations. Here, we instead use a SF recipe based on the H_2 content of the galaxy, motivated by recent observational results.

Bigiel et al. (2008) found, based on observations of spiral galaxies from the H I Nearby Galaxy Survey (THINGS), that the SF time-scale in molecular gas is approximately constant, i.e.

$$\Sigma_{\rm SFR} = A_{\rm SF}\,\Sigma_{\rm H_2}{}^{N_{\rm SF}} \quad (1)$$

with $N_{\rm SF} \simeq 1$.

Observations of higher density environments, such as starbursts and high-redshift galaxies, suggest that above a critical H_2 surface density, the SF time-scale becomes a function of $\Sigma_{\rm H_2}$ such that the SF law steepens. Recent work in which a variable conversion factor between CO and H_2 is accounted for suggests that $N_{\rm SF} \simeq 2$ for high $\Sigma_{\rm H_2}$ (Narayanan et al. 2012). This steepening is also expected on theoretical grounds (Krumholz, McKee & Tumlinson 2009; Ostriker & Shetty 2011). In PST14 and B14, we adopted as our default model an H_2-based SF recipe of the form

$$\Sigma_{\rm SFR} = A_{\rm SF}\left[\frac{\Sigma_{\rm H_2}}{10\,{\rm M_\odot\ pc^{-2}}}\left(1+\frac{\Sigma_{\rm H_2}}{\Sigma_{\rm H_2,crit}}\right)^{N_{\rm SF}}\right]. \quad (2)$$

In SPT15, we found that this broken power-law SF recipe produces better agreement with observations of SFRs and stellar masses in high-redshift galaxies than equation (1), so we adopt it in all of the models presented in this work. We adopt $A_{\rm SF} = 6.0\times10^{-3}\,{\rm M_\odot\ yr^{-1}\ kpc^{-2}}$, $\Sigma_{\rm H_2,crit} = 70\,{\rm M_\odot\ pc^{-2}}$, and $N_{\rm SF} = 1.0$, consistent with the observational results discussed above. As shown in PST14, these choices produce good agreement with the observed atomic and molecular gas fractions in nearby galaxies.

2.4 Stellar population and dust extinction

We store an array of ages and metallicities for the stellar population in each galaxy, based on the SF and enrichment history computed by the model. At the specified output redshift, we convolve this grid with the simple stellar population models of Bruzual & Charlot (2003) to obtain the spectral energy distribution (SED) of the unattenuated starlight. We use the models based on the Padova 1994 isochrones with a Chabrier (2003) IMF. We account for absorption by intergalactic hydrogen along the line of sight using the model of Madau et al. (1996).

Our model for dust extinction is described in S12, where predictions for the observable properties of galaxies from the rest-UV to FIR and from $z \sim 0$–5, and comparisons with available observations, were also presented. Briefly, the dust model is based on the model proposed by Charlot & Fall (2000). As in their model, we consider extinction by two components, one due to the diffuse dust in the disc and another associated with the dense 'birth clouds' surrounding young star-forming regions. The face-on extinction optical depth $\tau_{V,0}$ of the diffuse dust is assumed to be proportional to the metallicity of the cold gas times the column density of the gas. A slab model is used to compute the inclination-dependent extinction.

Additionally, stars younger than an age $t_{\rm BC}$ are enshrouded in a cloud of dust with optical depth $\tau_{\rm BC,V} = \mu_{\rm BC}\,\tau_{V,0}$, where we treat $t_{\rm BC}$ and $\mu_{\rm BC}$ as free parameters. Finally, to extend the extinction estimate to other wavebands, we assume a starburst attenuation curve (Calzetti et al. 2000) for the diffuse dust component and a power-law extinction curve $A_\lambda \propto (\lambda/5500\,{\rm \AA})^n$, with $n = 0.7$, for the birth clouds (Charlot & Fall 2000).

2.5 Model variants

In B14, we presented results for several different combinations of our two gas partitioning recipes (the pressure-based BR recipe and the metallicity-dependent and UV radiation field dependent GK recipe), and two assumptions about the distribution of cold gas in discs (the 'standard disc' model with $f_j = 1$, and the 'extended disc' model with $f_j = 2.5$). We also showed there that none of the 'standard disc' models were able to reproduce the column density distribution or other basic properties of DLAS. Therefore, in this paper, we focus on the 'extended disc' models, BRj25 and GKj25 in the notation of B14. The properties of SFGs in the 'standard disc' models are very similar.

The SAMs contain a number of free parameters. These are kept fixed to the same values used in SPT15, PST14, and B13, which were found to reproduce fundamental galaxy properties at $z = 0$ in their baseline models. We note that SPT15 and PST14 only consider models corresponding to $f_j = 1$. A larger value of f_j leads to more extended discs and lower gas densities overall, less efficient formation of H_2 and less efficient SF. B14 discuss the effects of increasing f_j to 2.5 on the calibration quantities.

3 GALAXY SELECTION

In order to compare the properties of DLAS to SFGs selected via their stellar emission, we begin by selecting catalogues of each using the same criteria as observational surveys. First, we present relations between stellar mass, *H*-band magnitude, rest-frame *B*-band magnitude, and halo mass for our overall galaxy population in Section 3.1. We then discuss how we generate catalogues of SFGs using several colour selection techniques that have commonly been used in the literature, and compare with observed populations selected in the same manner. In identifying DLAS, we adopt the same approach as was presented in B14, which we summarize in Section 3.2. Our analysis is focused on redshift $z \sim 2$.

3.1 Emission-selected galaxies

The top row of Fig. 1 shows the conditional probability of selecting a galaxy at $1.7 < z < 2.3$ with a specific stellar mass as a function of observed *H*-band magnitude and rest-frame *B*-magnitude. In spite of the complex SF and chemical enrichment histories in our models, and the inclusion of dust extinction, stellar mass is tightly correlated with both observed *H*-band and rest-frame *B* magnitude over the entire range in stellar mass. Similar to the upper-left panel, the bottom-left panel shows the conditional probability distribution for selecting a galaxy with a given SFR at the plotted observed-frame *R*-band magnitude. For moderate to high SFRs, the *R*-band magnitude can vary over two magnitudes, and the SFR of a galaxy with $R = 25.5$ can easily differ by a factor of 5 or more. SFR does become more tightly correlated with *R*-band magnitude for fainter galaxies (e.g. $R > 28$), as these galaxies have not produced enough metals to form a substantial amount of dust. The bottom-right panel shows the relation between stellar mass and halo mass. This figure allows for a rough comparison of galaxies with a given observed *H*-band or rest-frame *B*-band magnitude to stellar mass, and subsequently to halo mass. They also indicate that at an *H*-band magnitude limit of $H = 28$, galaxies will have a minimum stellar mass of $\sim$2–3 $\times\ 10^7\ {\rm M_\odot}$ and correspondingly a minimum halo mass of $\sim$2–3 $\times\ 10^{10}\ {\rm M_\odot}$. As *R*-band limited catalogues may miss galaxies with significant dust extinction, we focus the majority of our analysis on *H*-band or stellar mass limited samples, and only utilize

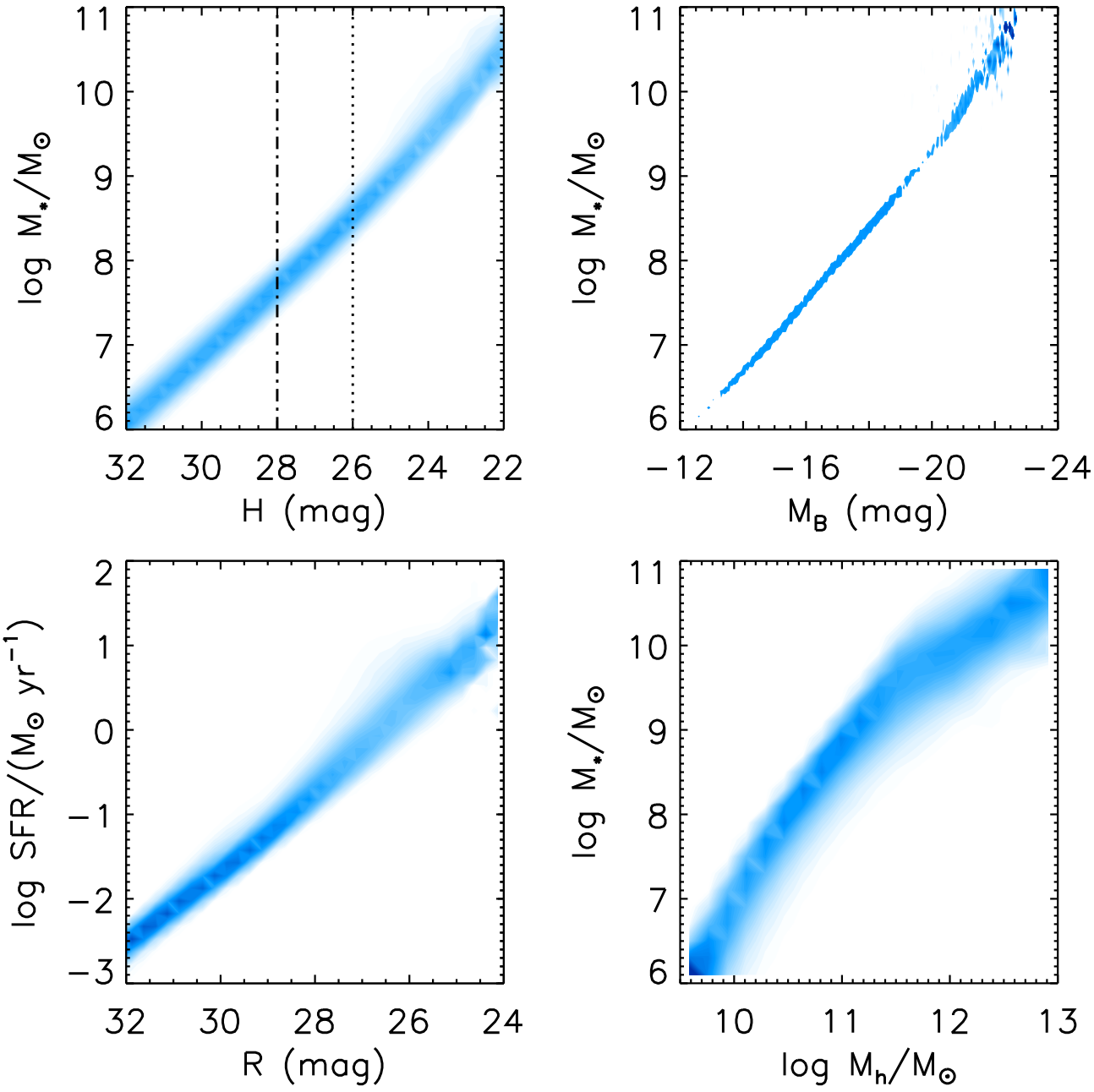

Figure 1. Top left – conditional probability distribution of stellar mass as a function of H-band magnitude for all galaxies at $1.7 < z < 2.3$ in the GKj25 model. Vertical lines show representative 'shallow' and 'deep' H-band limits (see text). Top right – same as top left except for stellar mass versus rest-frame B-band magnitude. Bottom left – SFR versus observed R-band magnitude. Bottom right – stellar mass versus halo mass. The first two panels illustrate how stellar mass tracks H-band magnitude and rest-frame B-band magnitude. The third panel shows the connection between SFR and observed-frame R-band magnitude.

the $R < 25.5$ limit as a subset when comparing to observational samples of SFGs with this magnitude limit.

In the past, various colour selection techniques have been used to identify high-redshift galaxies based on small numbers of photometric bands. We apply several colour selection criteria that have been commonly used in the literature to our models, in order to explore whether samples selected in this way are biased relative to underlying mass-selected samples and how they are related to DLAS.

Using the selection criteria of Steidel et al. (2004), which selects UV-bright, SFGs at $z \sim 2$, we identify a sample of BM/BX galaxies. These galaxies are selected using a colour–colour diagram in the $U - V$ and $V - R$ colours with an R-band magnitude limit of $R \leq 25.5$. These criteria take advantage of the Ly α decrement, similar to galaxies selected via the Lyman break, by selecting galaxies with blue rest-frame UV colours and absorption from intervening neutral hydrogen along the line of sight.

We also generate mock catalogues of *Hubble Space Telescope* (*HST*) Wide Field Camera 3 (WFC3) U_{225}, U_{275}, and U_{336} band dropouts using the $F225W$, $F275W$, $F336W$, B_{435}, and V_{606} filters and the selection criteria of Hathi et al. (2010). Similar to the BM/BXs and LBGs selected from ground-based telescopes, this technique takes advantage of the decrement in galaxy spectra blueward of Ly α. Hathi et al. (2010) find these galaxies to

have typical redshifts of $z = 1.6 \pm 0.2$, 2.1 ± 0.4, and 2.3 ± 0.4, respectively.

Massive galaxies at $1.4 < z < 2.5$ may be identified in K-selected samples based on their B-, z-, and K-band photometry (Daddi et al. 2004; Reddy et al. 2005). In this technique, requiring $BzK = (z - K)_{AB} - (B - z)_{AB} > -0.2$ selects SFGs (BzKs) at these redshifts, independently of their dust reddening. As star-forming BzKs exhibit a broad range of galaxy properties over a large redshift range, we focus our BzK analysis on only the brightest BzKs with K-band magnitudes $K < 23$ AB mag. Additionally, there is significant overlap between BXs, *HST* UV-dropouts, and BzKs, so this magnitude limit allows us to identify only the most massive galaxies. The relative insensitivity to dust reddening of this selection technique also permits us to probe SFGs with larger amounts of dust that could be missed from rest-UV colour selection methods.

We also 'observed' quiescent (non-star forming) galaxies and ultraluminous infrared galaxies. Quiescent galaxies were selected using BzK < -0.2 and $(z - K)_{\rm AB} > 2.5$. Ultraluminous infrared galaxies were selected based on their bolometric luminosities ($L_{\rm bol} > 10^{11}\,{\rm L}_{\odot}$). However, we found little overlap between these populations and DLAS, in the former due to a lack of a significant amount of cold gas and in the latter due to a low number density. Therefore we do not consider them further in this work.

The faint luminosities and small amounts of dust reddening in Ly α-emitting galaxies suggests that they may have similar physical properties to DLAS. Our SAMs do not track Ly α emission, which is extremely complicated to model in detail, so we are unable to identify galaxies that would be selected as Lyman alpha emitters (LAEs). None the less, Shapley et al. (2003) show that LAEs have similar rest-frame colours to LBGs and a large fraction would be selected as LBGs (and BM/BXs), if it were not for the $R \leq 25.5$ mag cut. Therefore, we do not compare the population of LAEs to DLAS, although we do investigate the connection between DLAS and fainter SFGs (e.g. BM/BXs), which may indirectly yield insight into the relation between DLAS and LAEs.

3.2 Damped Ly α absorption systems

For each galaxy, the molecular and atomic hydrogen gas is distributed in a disc with an exponential radial and vertical profile. We assume that the vertical scaleheight is proportional to the radial scalelength, $z_{\rm g} = \chi_z r_{\rm g}$, where $\chi_z = 0.4$ is a constant, in rough agreement with observations of moderate redshift galaxies. Varying the value of χ_z within reasonable limits (e.g. not razor-thin discs) has a minimal effect on our results.

The central gas density is then defined as $n_0 = M_{\rm cold}/(4\pi\mu m_{\rm H} r_{\rm g}^2 z_{\rm g})$, where $M_{\rm cold}$ is the mass of atomic and molecular gas, $m_{\rm H}$ is the mass of the hydrogen atom, and μ is the mean molecular weight of the gas. The atomic gas density as a function of radius and height above the plane is given by

$$n_{\rm H\,I}(r, z) = n_0 \left(1 - f_{\rm H_2}(r)\right) \exp\left(-\frac{r}{r_{\rm g}}\right) \exp\left(-\frac{|z|}{z_{\rm g}}\right). \quad (3)$$

See fig. 1 in B14 for examples of gas radial profiles for galaxies with different cold gas masses and metallicities in our models. The differences in the two gas partitioning recipes impact the gas distribution, amount of SF, and gas angular momentum, however, both recipes produce acceptable agreement with typical observed gas and SFR profiles in nearby galaxies (see also SPT15).

The models output the radial distance from the central galaxy for each satellite galaxy, and we assign a random azimuth and polar angle for each satellite's position with respect to the central. With the positions determined for every galaxy in each lightcone, we generate 20 000 random sightlines and integrate the three-dimensional gas density distribution along the sightline. Each galaxy is given a random orientation angle with respect to the sightline. All sightlines as well as the properties of all haloes with observed column densities above a threshold of $N_{\rm H\,I} > 10^{19}$ atoms cm^{-2} are then saved as our catalogue of absorption systems. In this chapter, we focus our analysis on systems with $N_{\rm H\,I} > 2 \times 10^{20}$ atoms cm^{-2}, i.e. DLAS.

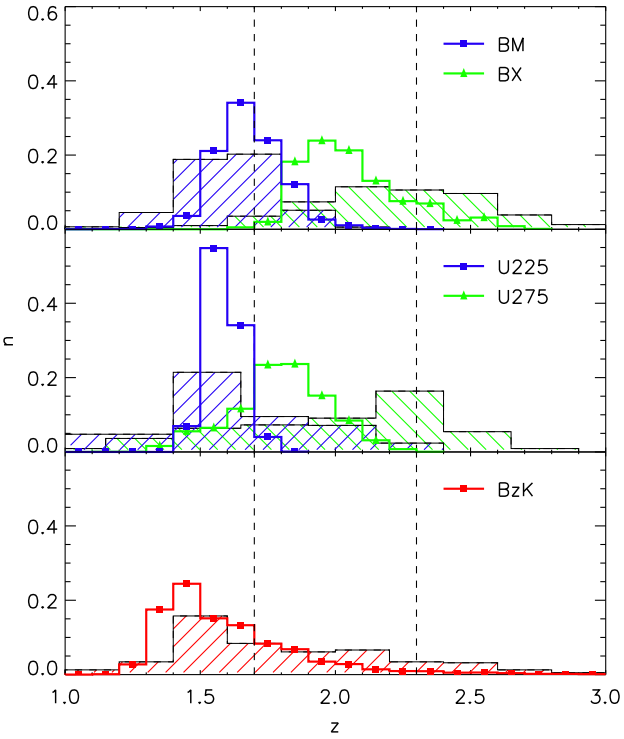

Figure 2. Redshift distributions of model and observed SFGs. Top panel – BM (blue squares) and BX (green triangles) galaxies in the GKj25 model with observations of BM and BX galaxies (blue and green hatched regions) from Shapley et al. (2005). Middle panel – same as top except for *HST* U225- (blue squares) and U275- (green triangles) band dropouts with observations from Hathi et al. (2010). Bottom panel – BzKs (red squares) with observations from Reddy et al. (2005).

4 PREDICTED AND OBSERVED PROPERTIES OF SFGs

In this section, we assess the accuracy of our models in reproducing the properties of SFGs by comparing those in our models to sets of observations. Fig. 2 shows the redshift distribution of SFGs identified in the GKj25 model and those observed in galaxy surveys. We find a similar redshift distribution in the BRj25 model (not shown). As can be seen in each panel, the distribution of model galaxy redshifts are qualitatively similar to observations. In all cases, we find a slightly narrower distribution of galaxy redshifts in the models than is observed. These differences are likely attributable to photometric errors in the observations. Table 1 shows the number densities of BM/BXs, U225-/U275-band dropouts, and BzKs for our models and observations. Both models underproduce the number density of SFGs, with the GKj25 model producing predictions closer to the observations.

As each galaxy population has a different redshift distribution and their properties vary slightly with redshift, we identify a subset with

Table 1. Galaxy properties.

Galaxy type	Selection limit	$\rho_{\rm obs}$ (arcmin^{-2})	$\rho_{\rm BR}$ (arcmin^{-2})	$\rho_{\rm GK}$ (arcmin^{-2})
BM/BX	$R < 25.5$	3.8/5.2	0.65/0.73	0.84/1.50
BzK	$K_s < 21$	3.1 ± 0.2	0.79	1.93
U225/U275	$B_{435} < 26.5$	–/–	0.49/0.47	2.54/3.67

redshifts $1.7 < z < 2.3$ in order to make a more robust comparison to DLAS at the same redshift. At this redshift, BM and BX galaxies exhibit similar properties, and therefore we combine them to form a catalogue of UV-bright SFGs. Similarly, we combine the *HST* U225- and U275-band dropouts into a catalogue of *HST* UV-bright, SFGs. In both cases, the majority of galaxies are the higher redshift BX galaxies and *HST* U275-band dropouts.

Fig. 3 shows the distribution of stellar masses, SFRs, and $E(B-V)$ values for observed and model galaxies in our BRj25 and GKj25 models at $1.7 < z < 2.3$. Overall, both models produce similar galaxy properties, which are again in qualitative agreement with observations. Similar to the distribution of observed galaxy redshifts, photometric errors likely contribute to the uncertainties in galaxy physical properties estimated via SED fitting. In comparison to observations, our models produce galaxies with a narrower range in SFRs and lower SFRs overall, mainly lacking the tail to high SFRs. Reproducing the very high SFRs observed in high-redshift galaxies is currently challenging for cosmological models (e.g. Niemi et al. 2012). Our models also produce larger $E(B-V)$ values and a broader distribution. For each class of SFG, the BRj25 model predicts higher stellar masses than the GKj25 model, varying from ∼0.2 to ∼1 dex. The BRj25 model also predicts higher SFRs in BM/BX galaxies and U225-/U275-band dropouts than the GKj25 model. In the remainder of the paper, due to the similarities in galaxy properties and redshifts for our two models, we show the properties of SFGs in the GKj25 model when the differences are not significant. Overall, in spite of producing too few galaxies with high SFR, the models produce galaxies with similar properties to observations.

5 RELATIONSHIP BETWEEN SFGs AND DLAS

In this section, we compare the predicted properties of colour-selected galaxies, DLAS, and a stellar mass-selected galaxy sample in our models. We begin by presenting the predicted $z = 2$ distribution functions for our models in Section 5.1. In Sections 5.2, 5.3, and 5.4 we present the properties and photometry of DLA host galaxies and compare them to galaxies at similar masses and luminosities. The differences between galaxy properties measured along the line of sight and the average for the host galaxy are presented in Sections 5.5 and 5.6. Finally, we discuss DLA host galaxies and SFGs in the context of galaxy scaling relations in Section 5.7.

5.1 General galaxy population

As a reference point, we first present the $z = 2$ galaxy stellar mass functions (GSMFs), star formation rate functions (SFRFs), H I mass functions (HIMFs), H_2 mass functions (H2MFs), H-band magnitude functions, and metallicity functions in Fig. 4. Unless otherwise specified, the distributions shown are for the whole 'unbiased' population of galaxies predicted by our models with no additional selection criteria applied. The upper-left panel of Fig. 4 shows the GSMFs at $z = 2$ for our models and observations. The predicted GSMFs for both of our models are in reasonable agreement with the observed GSMFs, although the models produce too many galaxies at low stellar masses. The overproduction of low-mass galaxies is a common problem for current cosmological models of galaxy formation (see Fontanot et al. 2009; Somerville & Davé 2015). Possible schematic solutions are presented in Henriques et al. (2013) and White, Somerville & Ferguson (2015). For both the GSMFs and SFRFs, we introduce uncertainty in measuring the SFRs and stellar masses by randomly taking a stellar mass and SFR from a distribution centred on the true values and with a conservative width of $\Delta M_{\rm s} = 0.2$ and $\Delta{\rm SFR} = 0.2$. This process has a negligible effect on the low-mass end of the distributions, but, as is well known, flattens the high-mass end. Similar to the upper-left panel, the upper-right panel of Fig. 4 shows the SFRFs for the models along with observational estimates. Again, our model predictions are in reasonable agreement with observations, although they underproduce the number of galaxies with high SFRs. While the BRj25 model is more successful at reproducing the GSMF, the GKj25 model does a better job with the SFRF. This result implies that the stellar mass to SFR ratio is not in agreement with observations, which indeed we shall later see is the case in Fig. 13. We note that although the BRj25 and GKj25 models produce results that are very similar, the differences seen here are larger than those shown in SPT15. This is because the extended disc model ($f_j = 2.5$) is more sensitive to the gas partitioning recipe than the standard disc model ($f_j = 1$) used in SPT15.

The middle left and right panels show the HIMFs and H2MFs, respectively, for both our models. The HIMFs in each of our models are similar at high masses although the GKj25 model produces fewer galaxies with intermediate H I masses and more galaxies with low-H I masses. Both models predict more H I in galaxies at $z = 2$ than is observed in local galaxies (not shown), although the shapes of the HIMFs are similar. The HIMFs for the $f_j = 2.5$ models do change more with redshift than the $f_j = 1$ models (e.g. see PST14). The GKj25 model predicts that more cold gas is in the molecular phase in galaxies of all masses, especially in more massive galaxies. This result is also reflected in the higher SFRs in the GKj25 model. Both models predict significantly flatter slopes than the local H2MF, which is mainly caused by a larger number of high-H_2 mass galaxies.

The bottom-left panel shows the observed frame H-band magnitude function for each model. We adopt representative 'shallow' and 'deep' magnitude limits of 26th and 28th magnitude in the remainder of this paper, corresponding to ∼2–3 orbits in the $H160W$ filter with WFC3 for the shallow limit and ∼10 orbits for the 'deep' limit (based on extended sources typical of those detected at $z \sim 2$ in the CANDELS survey; Grogin et al. 2011). The H-band traces rest-frame optical light ($\lambda \sim 5500$ Å at $z = 2$), which is mainly determined by the galaxy stellar mass. The GKj25 model produces more galaxies at all H-band magnitudes than the BRj25 model, similar to the GSMFs. See S12 and SPT15 for a more comprehensive comparison of the predicted luminosity functions.

The bottom-right panel shows the distribution functions of cold gas phase metallicity, where the metallicities are the mass-weighted average for each galaxy. The metallicity functions for the two models differ substantially at low metallicities, with the GKj25 model producing fewer galaxies with extremely low metallicities. We find it interesting that DLAS exhibit a metallicity floor of $[Z] \sim -2.6$ to -2.8, in line with the minimum metallicities observed in our models (Qian & Wasserburg 2003; Wolfe, Gawiser & Prochaska 2005; Rafelski et al. 2012).

In general, the $f_j = 1$ models (e.g. GKj1, BRj1; not shown) produce distribution functions quite similar to each other and to the

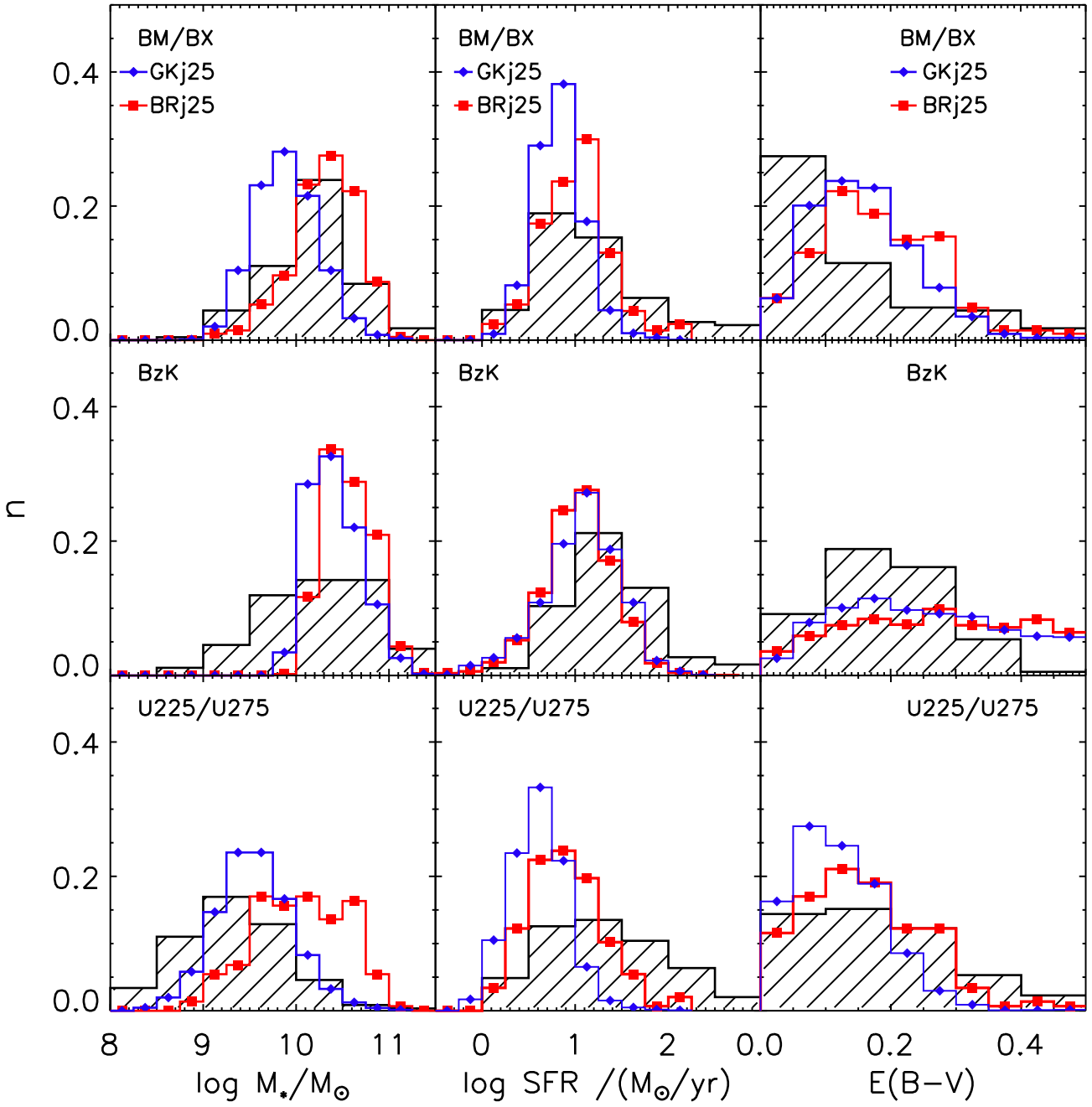

Figure 3. Normalized distribution of stellar masses (left-hand column), SFRs (middle column), and $E(B-V)$ values (right-hand column) are shown for the BRj25 model (red squares), GKj25 model (blue triangles), and observed galaxies (black) at $1.7 < z < 2.3$. References for observed galaxy samples are as in Fig. 2. Top row – BM/BXs. Middle row – BzKs. Bottom row – *HST* U225/U275 dropouts. In general, both models produce similar predicted galaxy properties, which typically have lower SFRs and comparable $E(B-V)$ values to observed SFGs. Stellar masses for model and observed BM/BXs and BzKs are comparable, yet the models produce higher stellar masses and lower SFRs than observed *HST* U225/U275 dropouts.

$f_j = 2.5$ models except they are more efficient at converting neutral gas into molecular gas, yielding higher stellar masses, SFRs, and metallicities. This difference is most pronounced at the upper end of the distribution functions. See SPT15 for further discussion of these models.

Of our two $f_j = 2.5$ models, both models are quite successful at reproducing the properties of SFGs, although the GKj25 model is our preferred model. B14 found both models to be fairly successful in reproducing the local GSMF, HIMF, and H2MF and the properties of high-redshift DLAS. Again, the GKj25 model was more successful overall. In Section 4, we showed that the properties of high-redshift SFGs do not vary significantly among our two models. Therefore, hereafter we focus our analysis on the GKj25 model when both models produce similar results.

5.2 DLA host galaxy properties

The DLA number densities for 10 000 sightlines at $1.7 < z < 2.3$ are overplotted at the bottom of each panel in Fig. 4. The normalization of these distribution functions is arbitrary, as the DLA cross-section

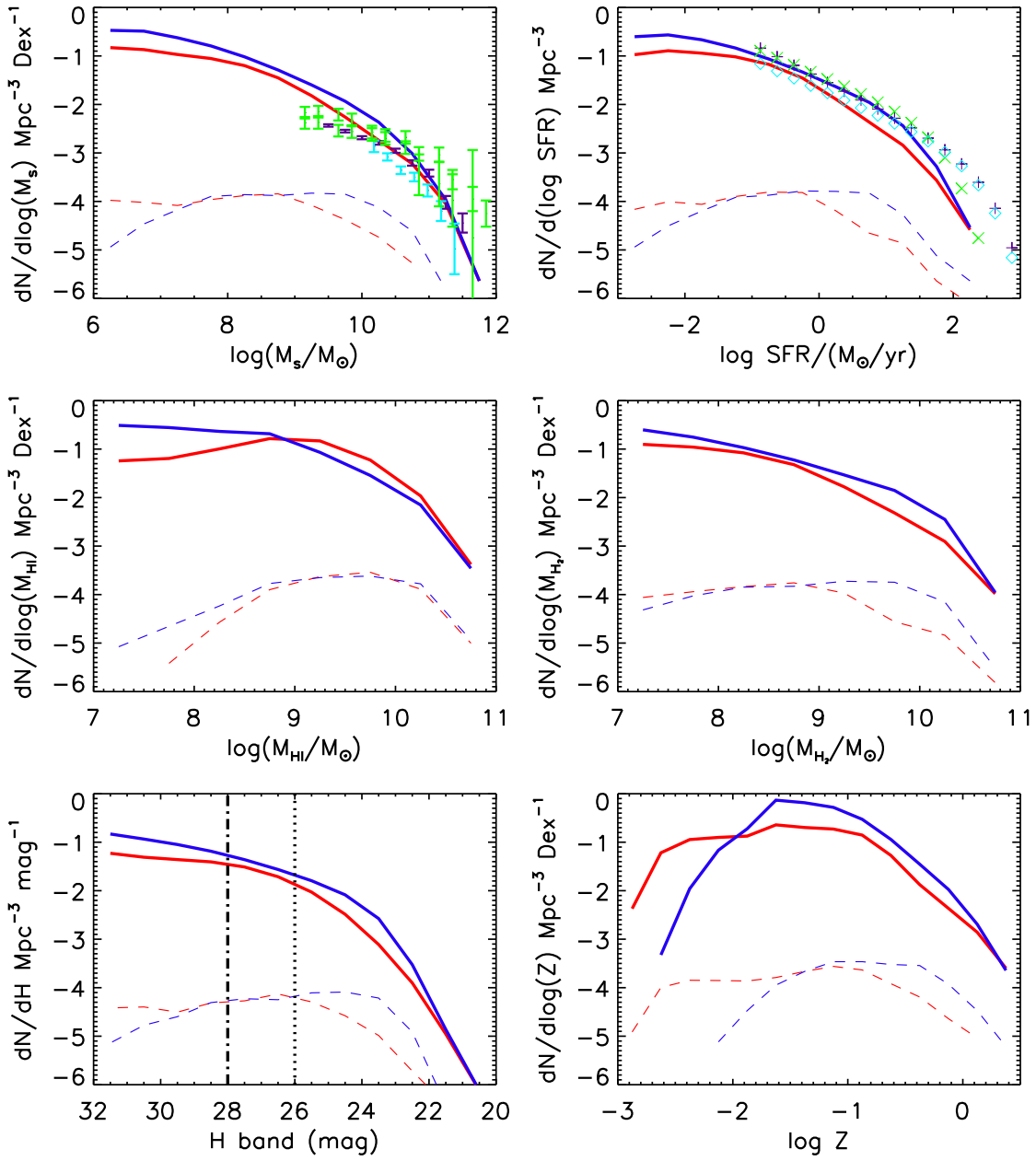

Figure 4. Top left – GSMFs (solid lines) and DLA stellar mass functions for 10 000 sightlines (dashed lines) in the BRj25 (red) and GKj25 (blue) models at $z = 2$. Top right – same as top left except for SFRFs. Middle left – HIMFs. Middle right – H2MFs. Bottom left – *H*-band magnitude functions. Bottom right – metallicity functions. Observations of GSMFs from Kajisawa et al. (2009, purple, $1.5 < z < 2.5$), Fontana et al. (2006, cyan, $1.6 < z < 2.0$), Marchesini et al. (2009, green, $1.5 < z < 2.0$ and $2.0 < z > 2.5$) are overplotted in the first panel. Observations of the SFRF from Reddy et al. (2008, purple, $z = 2.3$), Magnelli et al. (2011, cyan, $z = 2.0$), Sobral et al. (2012, green, $z = 2.2$) are also shown in the second. The representative WFC shallow (dotted line) and deep (dot–dashed line) *H*-band depths are also overplotted. Both models produce similar distribution functions, with the GKj25 model predicting slightly more efficient H_2 formation.

and number of sightlines are degenerate. However, comparing the shape of the distribution functions for DLAS and the overall galaxy population is quite interesting.

In terms of their stellar mass functions, DLAS in both models exhibit a relatively flat distribution, indicating that DLAS are almost equally likely to be selected across a large range in stellar mass, $6 \lesssim \log M_s/M_\odot \lesssim 10.5$. Similar to the GSMFs, there are fewer DLA host galaxies with large stellar masses and more with low stellar masses in the BRj25 model than in the GKj25 model. At the high-mass end, the shape of the DLA stellar mass distribution in both models tracks their respective GSMFs. Similar to the SMF, the *H*-band magnitude distribution function for DLAS is much flatter at the faint end than that of the underlying galaxy population. Moreover, our models predict that a substantial number of DLA host galaxies should be detectable with exposures of a few to 10 orbits with WFC3 on *HST*. We discuss the predicted detectability and photometric properties of DLAS further in Section 5.3.

Similar to the distribution of DLA stellar masses, DLAS in both models show a broad range in SFRs, $-3 \gtrsim \log$ SFR /M$_\odot$ yr$^{-1} \lesssim 1.5$, where DLAS in the BRj25 model typically have lower SFRs

Table 2. Average DLA host galaxy properties.

		GKj25	BRj25
$\log M_*$	$M_\odot$	8.7 ± 1.0	8.0 ± 1.1
log SFR	$M_\odot\ yr^{-1}$	-0.3 ± 1.0	-1.1 ± 1.0
$\log M_{H\,I}$	$M_\odot$	9.4 ± 0.7	9.5 ± 0.5
$\log M_{H_2}$	$M_\odot$	8.9 ± 0.9	8.1 ± 1.0
H (mag)	mag	26.2 ± 2.1	28.0 ± 2.5
$\log Z$	$Z_\odot$	-0.9 ± 0.4	-1.4 ± 0.7

Note. Errors show the 1σ confidence intervals.

than those in the GKj25 model. Unlike SFGs, the number density of DLAS flattens below SFR of $\sim$1 $M_\odot\ yr^{-1}$, and drops off below log SFR/$M_\odot\ yr^{-1} \lesssim -3$. Due to the dependence of SF on the density of molecular gas, we find a similar distribution of H_2 mass and SFR in both models. DLA host galaxies exhibit a broad range in H_2 mass which spans $\sim$4 decades.

Interestingly, due to the relatively flat distribution of stellar masses and SFRs over several decades, there is neither a well-defined 'typical' stellar mass and SFR nor a bimodality for DLAS. Due to the turnover at the high end of the GSMF and SFRF and the flat distribution of DLAS up to this turnover, our models predict that a large fraction of galaxies with high stellar masses and SFRs will host DLAS, specifically those at and above the knee in the stellar mass and SFR function. In both models, galaxies with stellar masses log $M_*/M_\odot \gtrsim 10$ and log SFR / ($M_\odot\ yr^{-1}$) $\gtrsim 1$ have the highest probability of hosting DLAS. Moreover, the number density of low-mass galaxies increases with decreasing stellar mass and SFR while the number of DLAS decreases slightly (these trends are slightly stronger in the GKj25 model than the BRj25 model). These trends imply that the fraction of low-mass galaxies hosting DLAS decreases below log $M_*/M_\odot \sim 10$ and log SFR/ ($M_\odot\ yr^{-1}$) ~ 1 in the GKj25 model. The BRj25 model shows similar trends. None the less in both models, although intermediate mass, SFGs are the most likely to host DLAS, they only make up a small fraction ($\lesssim$5–10 per cent) of the total number of DLAS, because low-mass galaxies are so much more numerous. The average DLA host galaxy properties for the GKj25 and BRj25 models at $z = 2$ are shown in Table 2.

As can be seen in the middle-left panel of Fig. 4, the most common H I mass for DLAS in both models is at log $M_{H\,I}/M_\odot \sim 9.5$, with the GKj25 model showing a tail to lower H I masses. Unsurprisingly, the H I mass is the property that is perhaps the best-defined characterization of DLAS in our models. Galaxies with the largest H I masses are the most likely to host DLAS. The broad distribution of H I masses is consistent with the large range of DLA cross-sections found in DLAS in simulations (Fumagalli et al. 2011; Cen 2012; B14).

The DLA metallicity function for the BRj25 model closely tracks the galaxy metallicity function, with DLA host galaxies found at all metallicities. The DLA metallicity function in the GKj25 model turns over at a metallicity of log $Z \sim -1$, while the BRj25 model has a flatter distribution that extends down to log $Z \sim -3$, our assumed pre-enriched metallicity. This result is interesting given that observed DLAS exhibit a metallicity floor at about log $Z \sim -2.5$ (Rafelski et al. 2012). We further explore the relation between the metallicities of DLAS and SFGs in Section 5.7. We note, however, that the location and slope of the low-metallicity turnover will be somewhat sensitive to resolution.

5.3 DLA photometry and overlap fraction

Fig. 5 shows the colours of all galaxies with $H < 28$ (left-hand column), the colours of DLAS with $H < 28$ (middle column), and the apparent magnitudes of DLAS at $1.7 < z < 2.3$ in the GKj25 model. DLAS and galaxies in the BRj25 model (not shown) have similar colour–colour diagrams, and the apparent magnitude distribution for the BRj25 model is also shown in the right-hand column. Note that DLA colours are the same as the typical galaxy colours at the same apparent magnitude. This indicates that DLAS do not select galaxies with any specific set of colours, nor should colour-selected galaxy samples be biased against DLA hosts. As the contours are dominated by the large number of faint galaxies, we overplot a brighter H-band limit of $H < 26$. Brighter DLAS are slightly redder in all colour–colour diagrams than fainter DLAS, similar to the galaxy population in general.

As can be seen in the right-hand column, both of our models predict that only a small fraction of DLAS are bright enough to be detected in optically selected surveys with relatively bright magnitude limits. This is due to the large number of DLAS with low stellar masses and SFRs. Both our models predict that probing down to $R = 28$ will result in a significant increase in the number of DLA counterparts over surveys limited to $R = 25.5$ at $z = 2$.

The left-hand panel of Fig. 6 shows the cumulative distribution for DLA hosts as a function of stellar mass at $z = 2$. Both models predict that the majority of DLA hosts have stellar masses greater than log $M_s/M_\odot > 8$. The cumulative distributions show that there are significantly more DLA hosts with high stellar masses in the GKj25 models. The right-hand panel of Fig. 6 shows the cumulative distribution of DLA hosts as a function of H-band magnitude. The GKj25 model predicts that 52 and 79 per cent of DLAS have H-band magnitudes $H < 26$ and 28, leading us to predict that a substantial fraction of DLA host galaxies should be detectable with WFC3 on *HST* in a moderate number of orbits. In contrast, the BRj25 model predicts that these detection rates drop to 24 and 55 per cent, respectively. Future observations that search for DLA hosts down to faint limits in the H band will help to discriminate between our two models for H_2 formation. Both of our models predict that there is a significant population of very faint DLA host galaxies that will require larger IR-sensitive telescopes such as the *James Webb Space Telescope* (*JWST*) to detect.

In both models, identifying DLAS with high metallicities (log $Z \gtrsim -1$) also selects objects with significantly larger stellar masses, and brighter H-band magnitudes (see Fig. 13). This is because DLAS, like the underlying galaxy population, have a positive correlation between stellar mass and metallicity, as discussed in more detail in Section 5.7. Therefore, searching for galaxy counterparts to high-metallicity DLAS should yield more detections at relatively bright magnitudes. Krogager et al. (2012) searched for DLA counterparts to high-metallicity DLAS which allowed them to more than double the number of detections at $z \gtrsim 2$ (also see Fynbo et al. 2010, 2011). The significant increase in detection rate is qualitatively more consistent with our BRj25 model, although we do note that the majority of DLAS in the GKj25 model have metallicities log $[Z] \gtrsim -1$.

5.4 Dust reddening

The top panel of Fig. 7 shows the distribution of $E(B-V)$ values for all DLA host galaxies and those with $H < 26$ in the GKj25 model. The bottom panel is similar except that is shows BzKs, BM/BXs, U225-/U275-band dropouts, and all galaxies

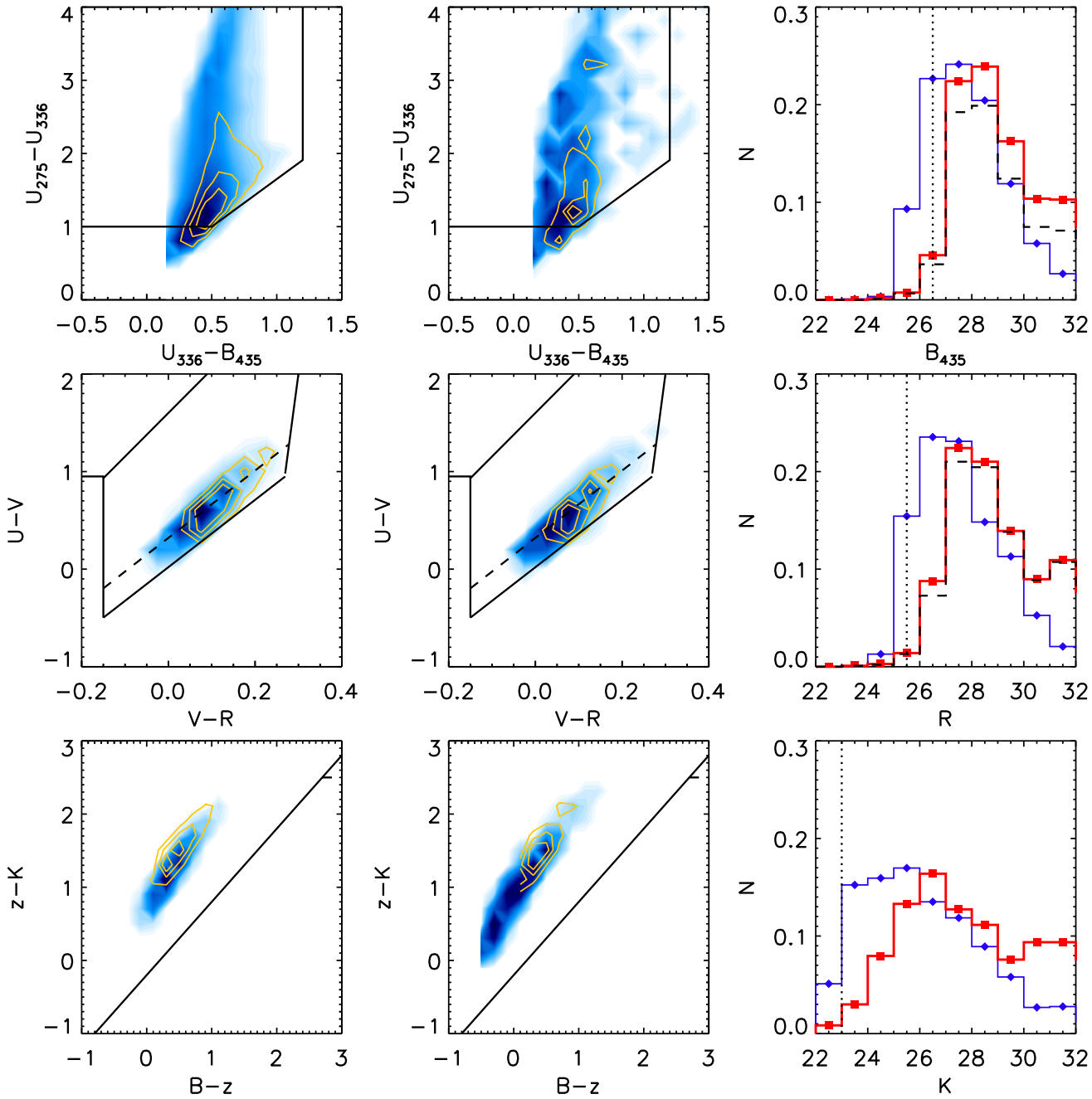

Figure 5. Left-hand column – colour–colour diagrams for all galaxies with $H < 28$ (blue) and $H < 26$ (orange contours) at $1.7 < z < 2.3$, with the colour selection regions used to identify high-redshift SFGs indicated. Middle column – same as left-hand column except for DLAS in the GKj25 model. Right-hand column – distribution of DLA B_{435}, R, and K-band magnitudes for DLAS in the BRj25 (red) and GKj25 (blue) models (top to bottom). From top to bottom, each row shows the colour–colour diagrams used to select *HST* U275-band dropouts, BM/BX galaxies, and BzK galaxies. The dashed line in the BM/BX colour–colour diagram shows the different selection region for BM (lower) and BX (upper) galaxies. The dashed lines in the right-hand column show the distribution of DLAS within the associated colour selection region for the BRj25 model. Vertical dotted lines indicate the typical magnitude limits for each of these surveys.

with $H < 26$. The distribution of $E(B-V)$ values is similar for the BRj25 model (not shown). These values reflect the average $E(B-V)$ value for each galaxy assuming a mixed dust geometry. In contrast, the results inferred from stacking quasars with and without DLAS along the line of sight as well as studies of DLAS towards radio-selected quasars reflect the average $E(B-V)$ value in DLA gas along the line of sight assuming a sheet geometry (e.g. Ellison et al. 2005; Khare et al. 2012). As the relation between these two quantities is non-trivial, we do not attempt to calculate the $E(B-V)$ value along the line of sight towards the DLA. None the less, we find that the vast majority of DLA host galaxies have very small $E(B-V)$ values, generally consistent with observations that suggest only trace amounts of reddening in DLAS (e.g. Murphy & Bernet 2016). For moderate to large impact parameters, the amount of dust reddening along the line of sight is likely lower than the average value for the host galaxy, while the effect of dust reddening will

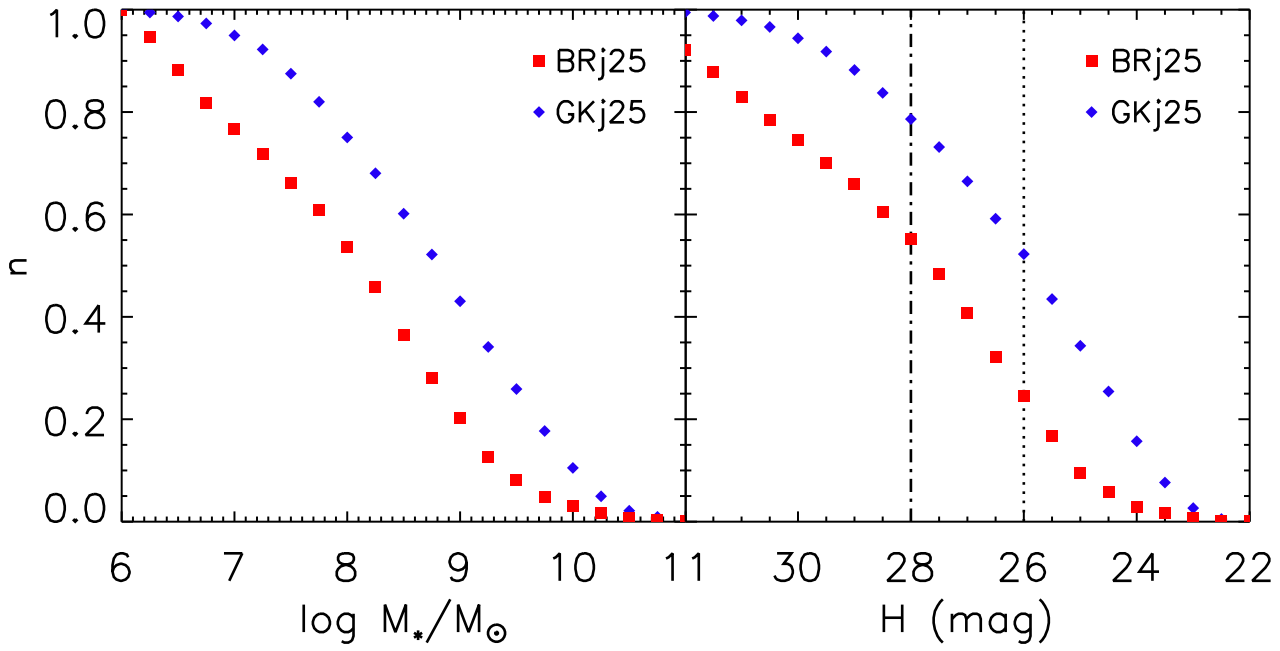

Figure 6. Left-hand panel – cumulative distribution of DLA host galaxy stellar masses in the BRj25 (red squares) and GKj25 (blue diamonds) at $z = 2$. Right-hand panel – same as left except for H-band magnitude. The vertical dotted and dashed lines at $H = 26$ and 28 show the representative 'shallow' and 'deep' observational limits. Our models predict that deep H-band imaging should yield a much higher fraction of detections of DLA hosts than previous optically selected studies. There are significant differences between the predictions of our two models, indicating that future observations of DLA host galaxies could begin to discriminate between different models for molecular hydrogen formation.

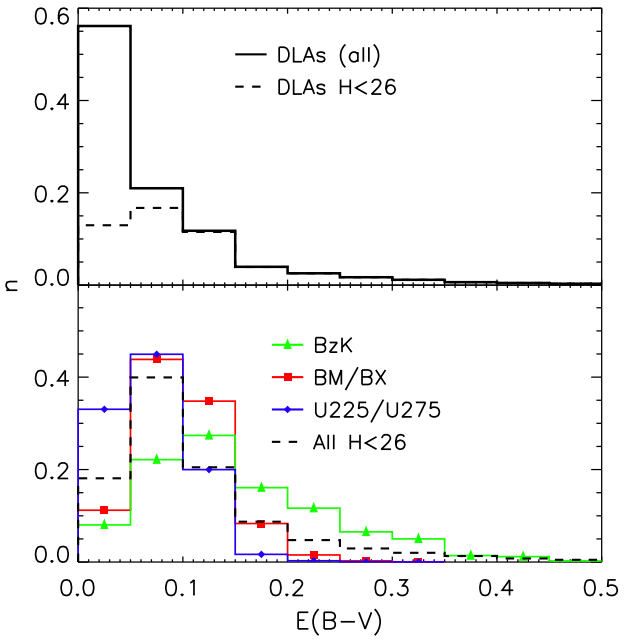

Figure 7. Upper panel – distribution of $E(B - V)$ values for all DLA host galaxies (black line) and those with H-band magnitudes $H < 26$. Lower panel – same as top except for BzKs (green triangles), BM/BXs (red squares), and *HST* U225-/U275-band dropouts (blue diamonds) with their associated magnitude limits as well as all galaxies with $H < 26$ (black dashed line). DLA host galaxies have very low $E(B - V)$ values, consistent with observations of DLAS. Model $E(B - V)$ values represent the average for each galaxy and assume a mixed dust geometry, while DLA $E(B - V)$ values likely act as more of a screen in front of the background quasar and only sample $E(B - V)$ along the line of sight. DLAS exhibit significantly smaller amounts of dust reddening than any type of SFG, as many of them are low-mass, metal-poor systems.

in general be stronger due to the difference in dust geometry. The majority of DLA host galaxies with low amounts of dust reddening are associated with low-mass, gas-rich haloes, while all DLA host galaxies with significant dust reddening, $E(B - V) > 0.1$, arise in galaxies with $H < 26$. Galaxies with $H < 26$ exhibit a tail to high $E(B - V)$, suggesting that these highly reddened DLAS may be missed from traditional colour-selected quasar surveys. They arise in the most massive galaxies with the highest metallicities, but represent only a small fraction of the DLA population. We predict the fraction of DLA host galaxies with $E(B - V) > 0.2$ to be 17 per cent in the GKj25 model and 7 per cent in the BRj25 model.

Galaxies identified based on their rest-frame ultraviolet colours (BM/BX galaxies and U225-/U275-band dropouts) have smaller amounts of dust reddening than the more massive BzKs, which are selected based on their rest-frame optical colours. Moreover, U225-/U275-band dropouts have less dust reddening than BM/BX galaxies due to the fainter magnitude limits of the available surveys. When we decrease the luminosity threshold to $H < 26$ for BM/BXs, BzKs, and U225-/U275-band dropouts (not shown), we find that the colour selection criteria has no significant effect on the distribution of $E(B - V)$ values. However, there is a fairly strong correlation between stellar mass or SFR and dust extinction in our models, in agreement with observations.

5.5 DLA impact parameter

As DLAS probe gas at different galactocentric radii, they contain indirect information on the size of DLA host galaxies as well as how DLA and galaxy properties vary as a function of radius. Fig. 8 shows the predicted distribution of impact parameters for DLAS in the BRj25 and GKj25 models at $1.7 < z < 2.3$ as well as measurements from DLAS with known optical counterparts at $1.92 < z < 3.40$ (Krogager et al. 2012). The distribution of impact parameters

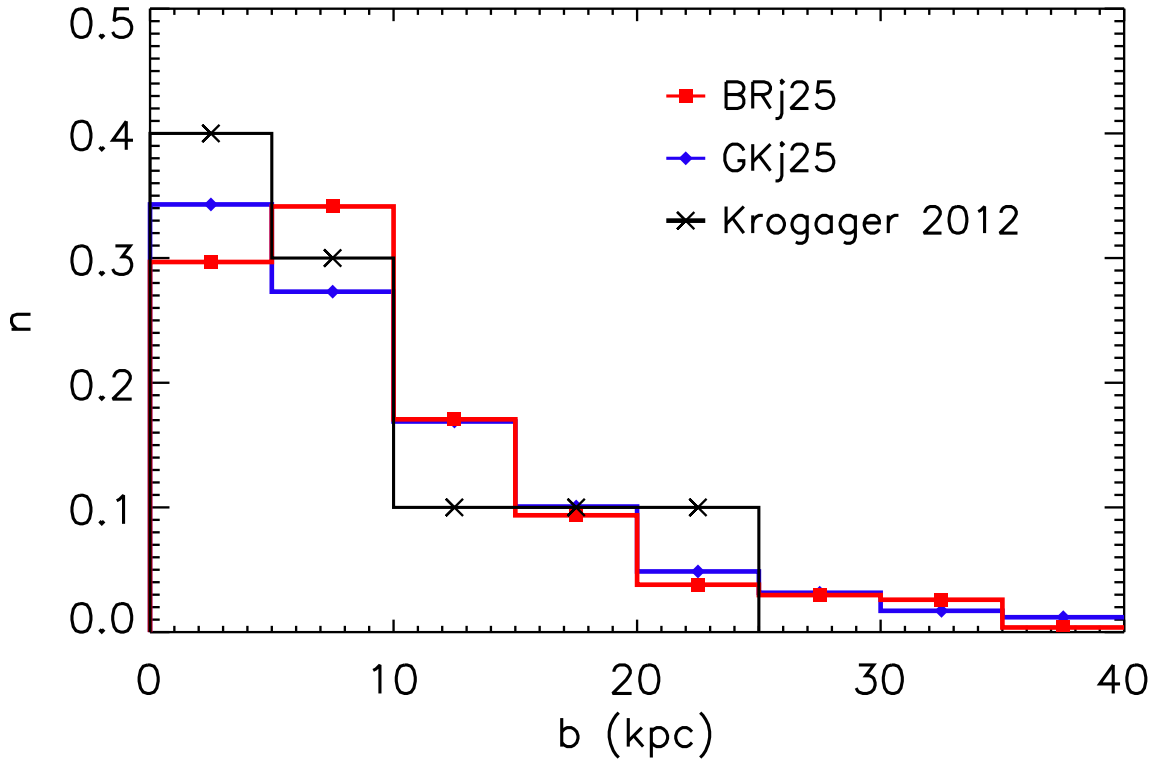

Figure 8. Distribution of impact parameters for DLAS at $1.7 < z < 2.3$ in the BRj25 (red squares) and GKj25 (blue diamonds) models, compared with observations of DLAS at $1.92 < z < 3.40$ (Krogager et al. 2012, black crosses). The distribution of DLA impact parameters DLAS in both models is in excellent agreement with observations. The majority of DLAS have impact parameters $b < 10$ kpc, or ∼1.2 arcsec.

in our models does not vary significantly from $z = 3$ to 2 (not shown). The predicted distribution of impact parameters in both models is remarkably similar to the observed one. For the models, we show the distribution of impact parameters from the DLA line of sight to the centre of the galaxy that is physically giving rise to the DLA. In the case where the line of sight passes through multiple galaxies, we show the impact parameter from the largest component. Observationally in the case where more than one potential counterpart is found, the natural preference is to measure the distance from the DLA to the nearest galaxy, which could bias the observed distribution to smaller values. With only 10 galaxy counterparts found, it is difficult to tell how significant this result is. Rahmati & Schaye (2014) examine the relation between different SFR thresholds and the identification of DLA counterparts in hydrodynamic simulations, and found that a significant number of DLAS originate in galaxies too faint to be detected in current studies, which could result in an association with a different, more luminous galaxy at a larger separation. Additionally, observations of DLAS may miss DLA host galaxies at very small impact parameters as the host galaxy may be difficult to identify against the glare of the bright background quasar. Krogager et al. (2012) targeted DLAS with metallicities $\log Z > -1$, which may have biased their DLAS to those hosted by more massive galaxies. When applying a similar metallicity cut (not shown), we find significantly more massive and more luminous DLA host galaxies. More massive galaxies have larger cold gas masses extending to larger radii, which would likely increase the average impact parameter.

As seen above, DLA sightlines often pass through the outer parts of galaxies, while emission-selected measurements are more likely to be weighted towards the denser central parts of galaxies. This could lead to a substantial offset between line-of-sight and global mass- or light-weighted average properties. We model the change in galaxy properties that would be measured for a DLA at its given impact parameter relative to the mass-weighted average properties of its host galaxy. Fig. 9 shows the distribution of H_2 fractions, SFRs, and metallicities for DLAS at $z = 2$ calculated along the DLA line of sight and averaged over the entire galaxy in the GKj25 model. We do not show the BRj25 model as the results are very similar. The H_2 fraction ($f_{H_2} = m_{H_2}(r)/(m_{H_2}(r) + m_{HI}(r))$) along the line of sight is calculated using the metallicity-based, varying-UV H_2 formation formulae and the cold gas density at the relevant radius. We use the average metallicity and SFR of the galaxy for simplicity. Due to the decreased gas density at the relatively large radius probed by typical DLA impact parameters, the H_2 fractions along the line of sight are significantly lower than the average H_2 fractions for each galaxy. Our models predict 85 per cent of DLAS have line-of-sight H_2 fractions $f_{H_2} < 0.1$, compared with 79 per cent for the mass-weighted average over the host galaxy. The SFR and metallicity both likely decrease with increasing radius, which would have the effect of decreasing and increasing the threshold for H_2 formation, respectively. The lower H_2 fraction and cold gas density also cause the SFR along the line of sight to be substantially lower than the average for the host galaxy. Our models predict that DLAS have typical line-of-sight SFRs ∼2–3 dex lower than the global average for their associated host galaxy. This may explain why some DLAS have such low SFRs yet can still be associated with relatively massive galaxies. In a small number of cases, the DLA line of sight passes through the centre of the galaxy where the cold gas density is *higher* than the average of the galaxy. Our models predict that these DLAS will exhibit larger H_2 fractions and higher SFRs.

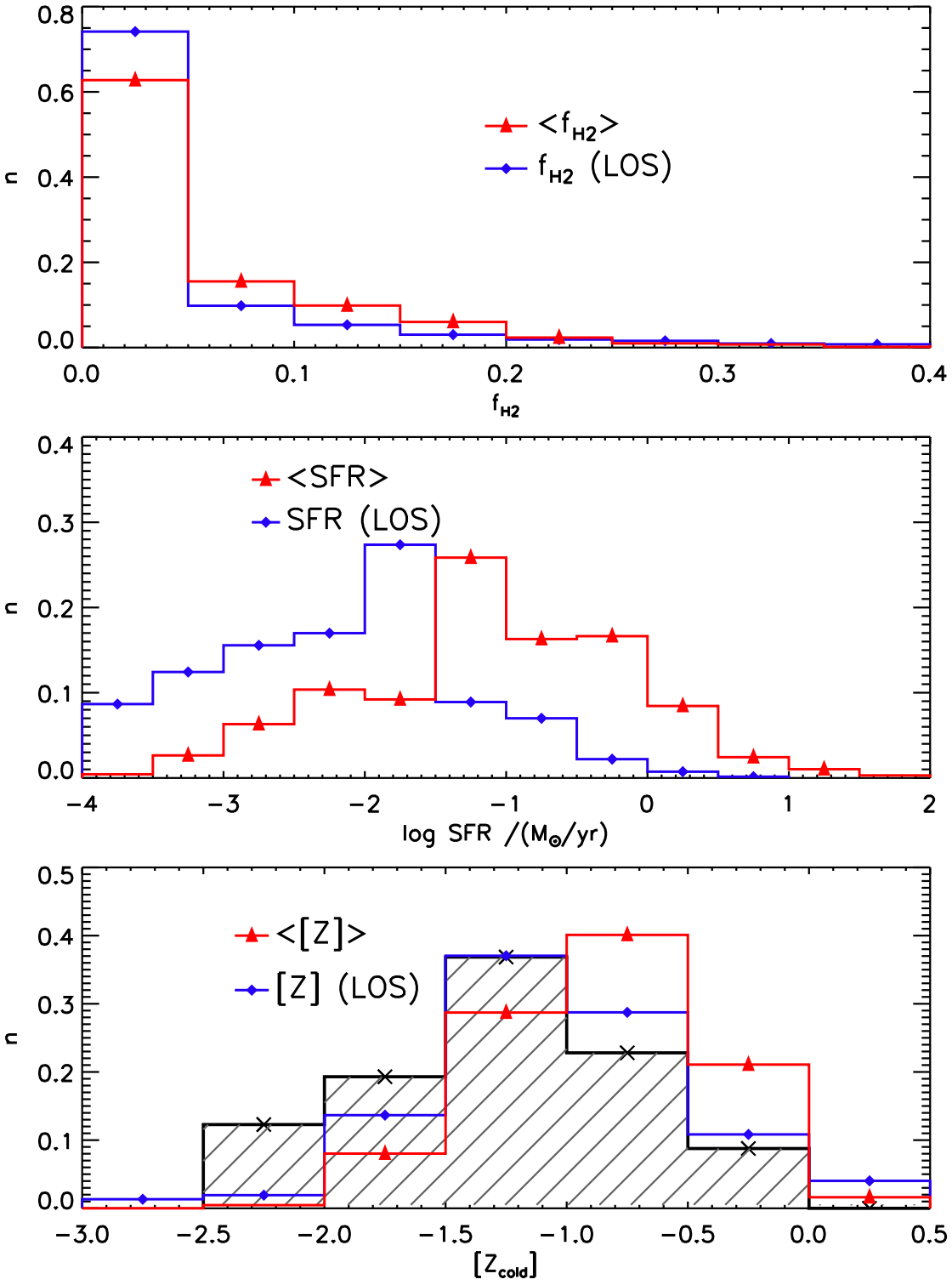

Figure 9. Top panel – distribution of average H_2 fraction in the DLA host galaxy (red triangles) and H_2 fraction along the DLA line of sight (blue diamonds) at $1.7 < z < 2.3$. Middle panel – same as top except for SFRs. Bottom panel – same as top except for metallicity, with an assumed metallicity gradient of $dZ/dr = -0.1$ dex kpc^{-1}. Observations of DLAS at the same redshift (Rafelski et al. 2012) are overplotted in black. All three quantities along the line of sight are substantially lower than the galaxy average, implying there may be a significant offset between quantities measured at the DLA line of sight and the mass-weighted average for the host galaxy.

5.6 Metallicity gradients

B14 found that modest metallicity gradients ($dZ/dr = -0.1$ dex kpc^{-1}), as observed in $z \sim 2$ SFGs, may contribute to the low metallicities observed in DLAS. Following this work, we further study the effect of metallicity gradients on the properties of $z \sim 2$ DLAS. Metallicities along the line of sight are calculated assuming a metallicity gradient of $dZ/dr = -0.1$ dex kpc^{-1} where the

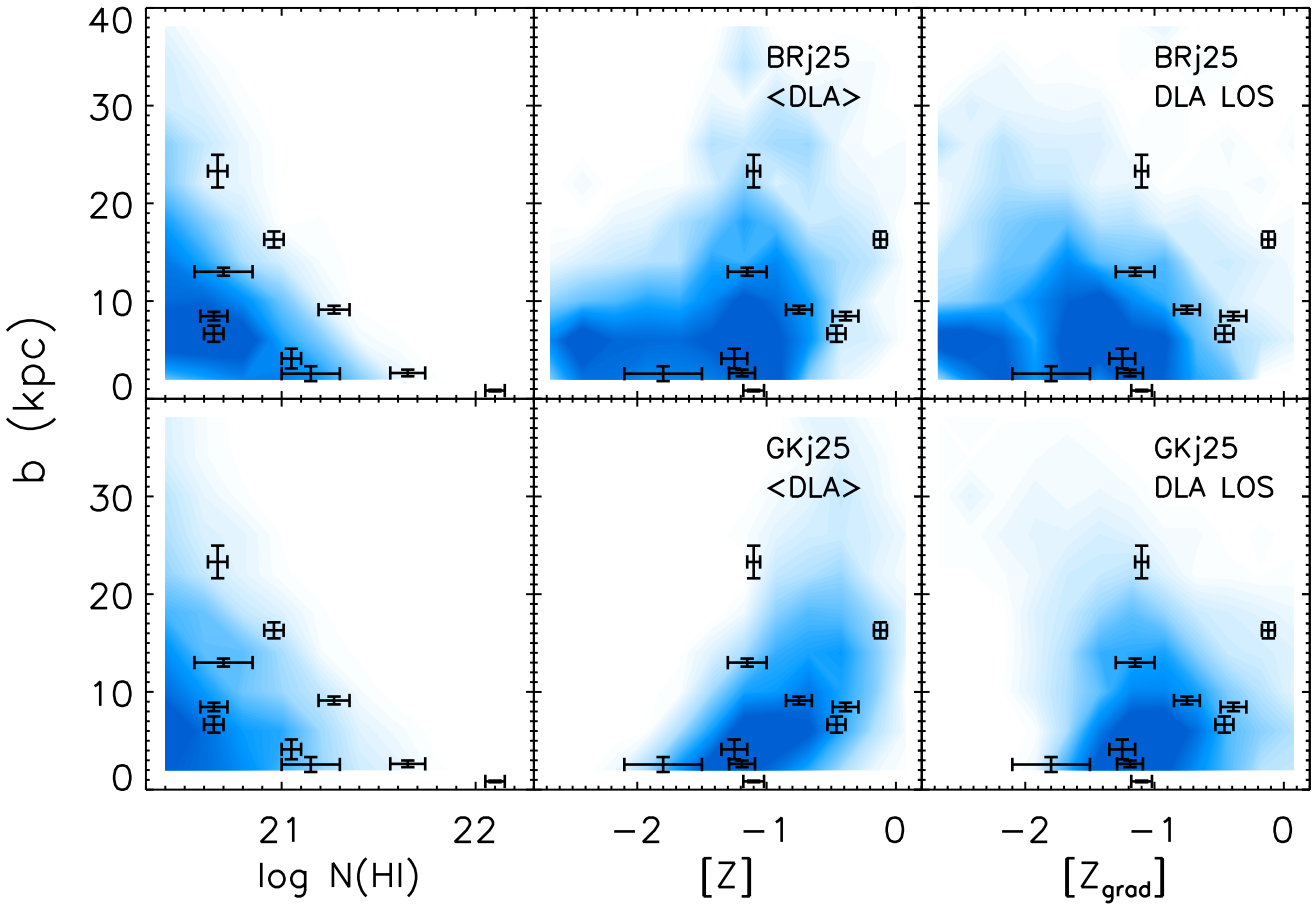

Figure 10. Probability of selecting a DLA in the BRj25 (top) and GKj25 (bottom) models at $1.7 < z < 2.3$ with a given impact parameter as a function of $N_{\rm H\,I}$ (left), average galaxy cold gas phase metallicity (middle), and metallicity along the line of sight including a metallicity gradient (right). Observations of DLAS at $1.92 < z < 3.40$ from Krogager et al. (2012) are overplotted. A clear trend is seen between larger impact parameter and both lower $N_{\rm H\,I}$ and lower metallicity. When a metallicity gradient is included, the trend predicted by the BRj25 model appears to be in conflict with the observations. However, the observed sample is quite small.

average metallicity is set at the radius of the average cold gas mass (see B14, for details). The inclusion of a modest metallicity gradient causes the model DLA metallicities to be more consistent with those of observed DLAS. DLA host galaxies with large stellar masses ($\log M_*/{\rm M}_\odot > 10$) have a similar metallicity distribution as SFGs of the same mass as in Kulas et al. (2013). In studies of individual DLAS and their host galaxies, Krogager et al. (2013) and Bouché et al. (2013) found that the DLA metallicity measured from interstellar absorption lines was ∼0.2–0.5 dex less than that inferred from nebular emission lines within the host galaxy.

The left-hand column of Fig. 10 shows the distribution of impact parameters as a function of $N_{\rm H\,I}$ for DLAS in both models at $z \sim 2$. Similar to observations, both models produce DLAS which have lower column densities at larger impact parameters, and high-$N_{\rm H\,I}$ DLAS that typically have smaller impact parameters. The models predict that there should be many objects at low impact parameters and low-column densities, which are not present in the observed sample. However, as lower column density DLAS are more likely to come from lower mass galaxies, detecting them in the proximity of a nearby quasar may be difficult. The relation between $N_{\rm H\,I}$ and impact parameter in our models is easily understood – as cold gas is distributed in exponential discs, high-$N_{\rm H\,I}$ DLAS will preferentially be selected at low impact parameters where cold gas densities are highest. The majority of low-$N_{\rm H\,I}$ DLAS are found in low-mass galaxies with small impact parameters, while DLAS at large impact parameters preferentially sample the outskirts of galaxies where $N_{\rm H\,I}$ values are lower.

The middle and right-hand columns of Fig. 10 show impact parameter as a function of DLA metallicity with and without an assumed metallicity gradient ($\mathrm{d}Z/\mathrm{d}r = -0.1$ dex kpc^{-1}), respectively. The general trend of more massive DLAS having higher metallicities underlies the predicted relation between impact parameter and metallicity. Observations indicate that galaxies with higher metallicities have larger cold gas masses extending to larger radii, making it more likely for them to be selected as DLAS at larger impact parameters. For the same reason, more massive metal-rich galaxies will be less likely to host DLAS at small impact parameters, explaining the lack of DLAS observed with low impact parameter and high metallicity. Although not included in our models, higher metallicities will also make it more likely for the central regions of these galaxies to contain more dust obscuration, which may redden or extinguish quasars, causing them to drop out of optically selected quasar surveys. Conversely, galaxies with low metallicities are less massive and only have high-$N_{\rm H\,I}$ gas at small radii. As the low metallicity, low impact parameter DLAS arise in low-mass galaxies, we expect observations of DLA host galaxies to miss this faint population.

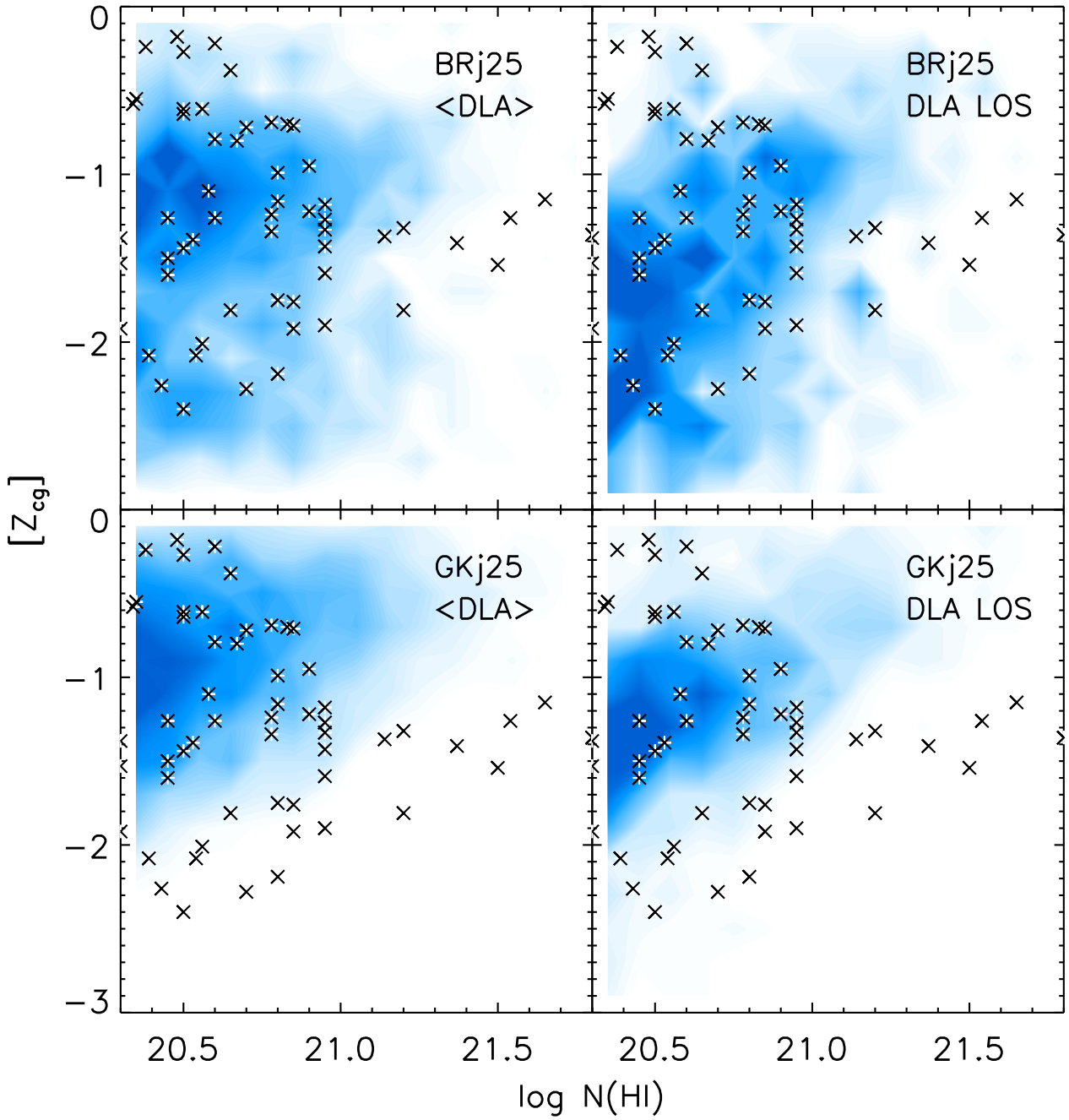

Figure 11. Probability of selecting a DLA with a given metallicity as a function of column density for the BRj25 model (upper row) and GKj25 (lower row) at $1.7 < z < 2.3$. DLA host galaxies (left) and DLA metallicity along the line of sight with a metallicity gradient (right). Observations of DLAS at $1.5 < z < 2.5$ from Rafelski et al. (2012) are overplotted. A metallicity gradient lowers the metallicity of low $N_{\rm H\,I}$ systems while having a small impact on higher metallicity ones. None of our models match the observations well in detail.

Given these trends, the BRj25 model produces a large number of DLAS with a large range of average metallicities and low impact parameters. Conversely, the GKj25 model is more successful at reproducing the observed trend between impact parameter and metallicity. Since metallicity gradients preferentially change DLA metallicity based on impact parameters, DLAS observed at high impact parameter have their metallicities change the most. In our models, we see this as a shift of DLAS horizontally to lower metallicities. The GKj25 model predictions are in plausible agreement with the observations, both with and without a metallicity gradient, while the BRj25 model does not reproduce the observed correlations, and produces results that are more obviously discrepant with the observations when a metallicity gradient is included. However, the existing sample with these kinds of measurements is extremely small, and may suffer from selection biases that we have not attempted to model in detail, so it is not possible to draw any strong conclusions.

Fig. 11 shows the distribution of DLA metallicity versus column density for both of our models with and without a metallicity gradient. Neither of our models reproduce the distribution well in detail

as they fail to reproduce a population of high-$N_{\rm H\,I}$ systems with low metallicities. The presence of these galaxies is hard to explain in the classical picture of DLAS as the highest column density gas typically resides in the centres of galaxies where metallicities are higher. Since metallicity gradients have the largest effect on DLAS at larger impact parameters where column densities are lower, they are unable to explain these results. An alternative explanation is that these systems may be 'contaminated' by pristine gas such as in intergalactic streams or clumps, which may be increasing the H I column density and decreasing the metallicity. The BRj25 model predicts a larger range in DLA metallicities than the GKj25 model. The GKj25 model also shows more of a trend between DLA metallicity and column density. Both of our models predict the existence of high column density systems with higher metallicities, which are not seen in observations. This may be a selection effect as they likely also have more dust reddening due to their high column density and metallicity.

5.7 Galaxy scaling relations

In this section, we compare scaling relations for model galaxies selected in different ways: (1) as DLAS, (2) as SFGs, using a colour-based selection technique and an apparent magnitude limit or (3) a hypothetical 'unbiased' stellar-mass selected sample. In this section we identify SFGs using the BM/BX selection criteria, as we find that applying the BzK or *HST* U225/U275 selection criteria produces similar results. Our main SFG sample is magnitude limited in the observed H band. We also show results for SFG samples with a brighter, observed R-band limit. We only show the GKj25 model as the BRj25 model exhibits similar trends. The shaded regions in the upper row of Fig. 12 show the conditional probability of selecting a DLA host galaxy (left-hand panel) or a SFG (right-hand panel) with a given DLA cross-section as a function of stellar mass in the GKj25 model. A galaxy's DLA cross-section, $\sigma_{\rm DLA}$, is defined as the inclination corrected cross-sectional area in kpc^2 for which its H I column density is $N_{\rm H\,I} > 2 \times 10^{20}$ atoms cm^{-2}. The mean and $\pm 1\sigma$ of the conditional probability distribution for an unbiased stellar mass selected sample is also shown.

DLAS preferentially select galaxies with higher $\sigma_{\rm DLA}$ values by $\sim$0.5 dex than a typical galaxy at a given stellar mass. This difference is larger at lower stellar masses. This trend results in a flatter slope for the $\sigma_{\rm DLA}$–M_* relation. As DLAS are selected based on their DLA cross-section, a smaller fraction of low-mass galaxies will have sufficiently large $\sigma_{\rm DLA}$ values to be identified as DLAS. For the most part, SFGs follow the $\sigma_{\rm DLA}$–M_* relation seen in the overall (stellar mass selected) galaxy population with a slightly higher DLA cross-section at a given stellar mass. SFGs with $R < 25.5$ select galaxies with systematically higher $\sigma_{\rm DLA}$. The magnitude limit of $R < 25.5$ effectively picks out galaxies with an SFR threshold, which translates approximately into a minimum cold gas mass.

The lower row of Fig. 12 is similar to the upper row except that the conditional probability distribution of the cold gas fraction ($f_{\rm g} \equiv \frac{M_{\rm H\,I}+M_{\rm H_2}}{M_{\rm H\,I}+M_{\rm H_2}+M_{\rm s}}$) is shown as a function of stellar mass. DLAS and colour-selected SFGs follow the general trend of decreasing cold gas fraction with increasing stellar mass, as is seen in the stellar mass selected sample, while both SFGs and DLAS preferentially select galaxies with larger amounts of cold gas. At low stellar masses, DLAS are almost entirely composed of cold gas. SFGs show a similar trend, though it is not as extreme. Low stellar mass SFGs with $R < 25.5$ have significantly higher cold gas fractions than the underlying population at that mass, again because of the correlation between SF activity and cold gas fraction that is built into our models.

Similar to Fig. 12, the upper row of Fig. 13 shows the conditional probability distribution for SFR at a given stellar mass for DLA host galaxies and colour-selected SFGs. The models produce galaxies with SFRs that are lower than the observations by $\sim$0.5 dex at a given stellar mass. This is a common and well-known problem in cosmological models (see Section 6). For the most part, both DLAS and SFGs follow the SFR–$M_{\rm s}$ relation of the stellar mass selected galaxy population. DLA host galaxies produce slightly higher SFRs at low stellar masses, again because of the bias towards high gas fractions. For low-mass SFGs, the $R < 25.5$ mag limit again selects galaxies with higher SFRs, leading to a slightly flatter SFR–$M_{\rm s}$ relation.

The mass–metallicity relation for DLAS with and without a metallicity gradient ($\mathrm{d}Z/\mathrm{d}r = -0.1$ dex kpc^{-1}) and SFGs is shown in the lower row of Fig. 13. The stellar mass selected population of model galaxies reproduces the observed mass–metallicity relation quite well. Both DLA hosts without a metallicity gradient and SFGs show a trend with stellar mass similar to typical galaxies, but offset to slightly lower metallicities. An even larger offset is seen for SFGs with $R < 25.5$. This is expected, based on the 'Fundamental Mass–Metallicity Relation' which implies that galaxies with higher SFR (and higher gas fractions) have lower metallicities at a given stellar mass (Mannucci et al. 2010). When a metallicity gradient is included, DLAS generally show an increase in metallicity with stellar mass although with considerably more scatter. This is in part due to how we model the metallicity gradient as larger galaxies will have a larger range in impact parameter, yielding a larger range in metallicity shifts. In the case of a metallicity gradient in high-mass galaxies, this large range in metallicity shifts would have the effect of washing out the mass–metallicity relation for DLAS at $\log M_{\rm s}/{\rm M}_\odot \gtrsim 9$. More extreme metallicity gradients or a range of metallicity gradients would further contribute to the scatter. Therefore, the absence of a mass–metallicity relation in high-mass DLAS may indicate that metallicities vary substantially with galactocentric radii.

Fig. 14 is similar to Fig. 12 except that $E(B-V)$ values are shown. The higher metallicities present in more massive galaxies cause the amount of dust reddening to also increase substantially with stellar mass. This trend is also seen in DLA host and SFGs, although both techniques select galaxies with lower amounts of dust reddening. For DLA host galaxies, this is due to their large cold gas masses selecting galaxies with higher specific SFRs and therefore lower average metallicities. A similar selection effect applies to SFGs as they have preferentially higher SFRs and blue colours. Our model predictions for $E(B-V)$ are generally consistent with observationally derived values for galaxies with the same magnitude limit and redshift, although observations do exhibit considerable scatter in the distribution of $E(B-V)$ values. This may be due to photometric errors or the relatively simple approach we have used to model dust extinction. It could also be due to differences in the dispersion of gas or metal content between the models and observations. In the models, galaxies with stellar masses below $\log M_{\rm s}/{\rm M}_\odot < 8$ have very small amounts of dust reddening due to their low metallicities. These results suggest that the highest mass DLAS may have substantial amounts of reddening and would be the most likely to be missed in colour-selected quasar surveys. However, according to our models, they form only a small fraction of the total number of DLAS.

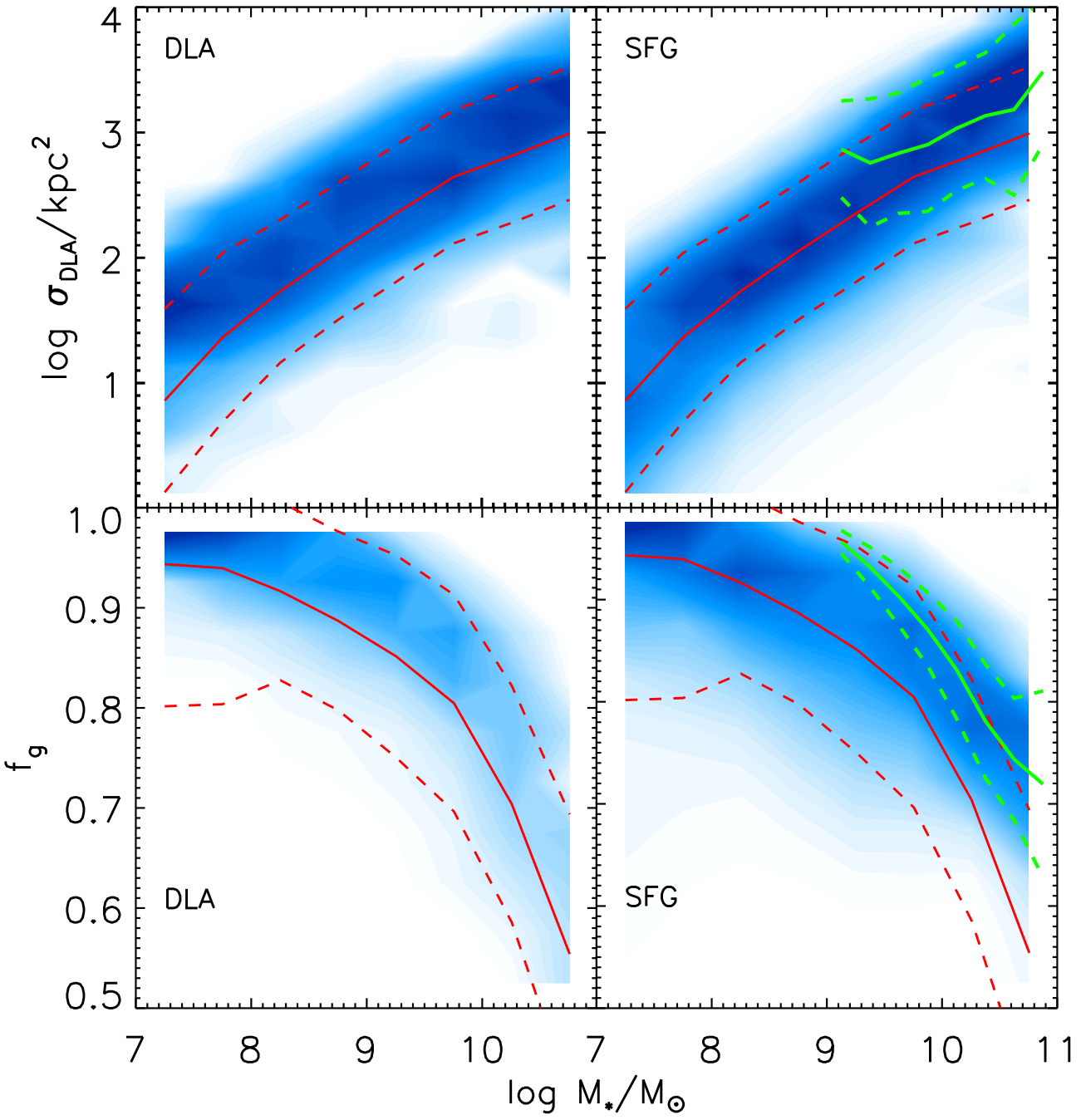

Figure 12. Upper row: blue shaded regions show the conditional probability of selecting a galaxy with a given DLA cross-section ($\sigma_{\rm DLA}$) as a function of stellar mass, for DLAS (left) and SFGs identified using the BM/BX selection criteria (right) in the GKj25 model. The average and 1σ range of the conditional probability for an unbiased stellar mass selected sample is shown via the red solid and dashed lines. The green lines show the mean and 1σ range for a magnitude-limited SFG sample ($R < 25.5$). Lower row: same as top except for the conditional probability of the cold gas fraction. Both (luminous) SFGs and DLAS preferentially select galaxies with higher $\sigma_{\rm DLA}$ and $f_{\rm g}$ values at a given stellar mass than an 'unbiased' stellar mass selected sample. In the case of the SFGs, this is primarily due to the assumed magnitude limit, which effectively selects galaxies with a minimum SFR and therefore a minimum cold gas mass.

6 DISCUSSION

Studies of atomic hydrogen at high redshift to date have been limited to objects selected in absorption against background quasars. In particular, DLAS have traditionally been thought to contain a substantial fraction of the atomic hydrogen in the universe, and therefore were considered important reservoirs for future SF. However, the connection between DLAS and galaxies identified via their stellar emission remains unclear. In this paper, we made use of cosmological models of galaxy formation to make explicit and detailed predictions about the nature of this relationship. Our models allow us to examine which types of galaxies are selected as DLAS, how DLAS relate to SFGs identified through common selection techniques, and how line-of-sight DLA properties differ from the mass-weighted averages for their host galaxies. We considered two different models for the formation of molecular hydrogen that we presented in previous works (PST14, B14, SPT15): a pressure based

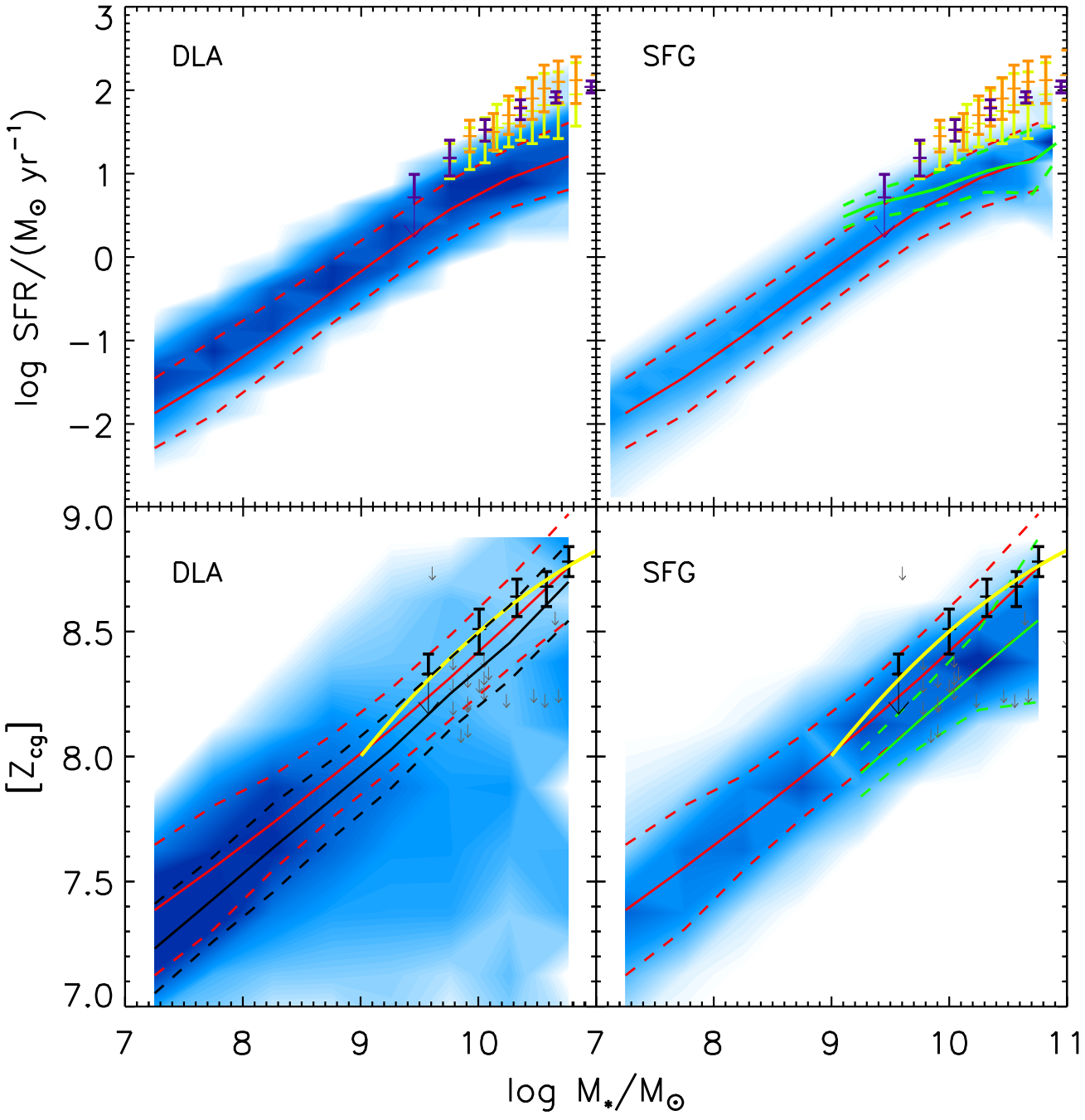

Figure 13. Same as Fig. 12 except that the conditional probability of SFR (upper row) and DLA metallicity (lower row) for a given stellar mass is shown. Observational estimates from Karim et al. (2011, purple) at $1.6 < z < 2.5$ and from Whitaker et al. (2011, yellow, $1.5 < z < 2.0$; orange, $2.0 < z < 2.5$) are overplotted in the upper row. Average metallicities for SFGs from Erb et al. (2006, black bars), individual BX galaxy metallicities from the field sample of Kulas et al. (2013, grey arrows), and the best-fitting line to Erb et al. (2006) from Maiolino et al. (2008, yellow) are shown in the lower row. The stellar mass selected galaxy population agrees well with the Erb et al. (2006) results, while SFGs are offset to lower metallicities at a given stellar mass. The average DLA host galaxy metallicity without a metallicity gradient is overplotted in black, while the shaded areas include a metallicity gradient in the modelling. Applying a metallicity gradient in DLAS effectively washes out the mass–metallicity relation at higher stellar masses.

model that we refer to as the BRj25 model and a metallicity based approach referred to as GKj25.

6.1 Properties of DLA host galaxies

One interesting insight is that both of our models predict that the number density of $z \sim 2$ DLA host galaxies at a given stellar mass or SFR is constant over several decades. This suggests that massive galaxies are more likely to host DLAS, but DLAS have a relatively equal probability of originating from galaxies with a large range of masses. Therefore, DLA host galaxies are characterized by a broad distribution of properties. Since a large number of DLA host galaxies arise in relatively low-mass haloes, our models predict that many are too faint to be probed by current emission-selected

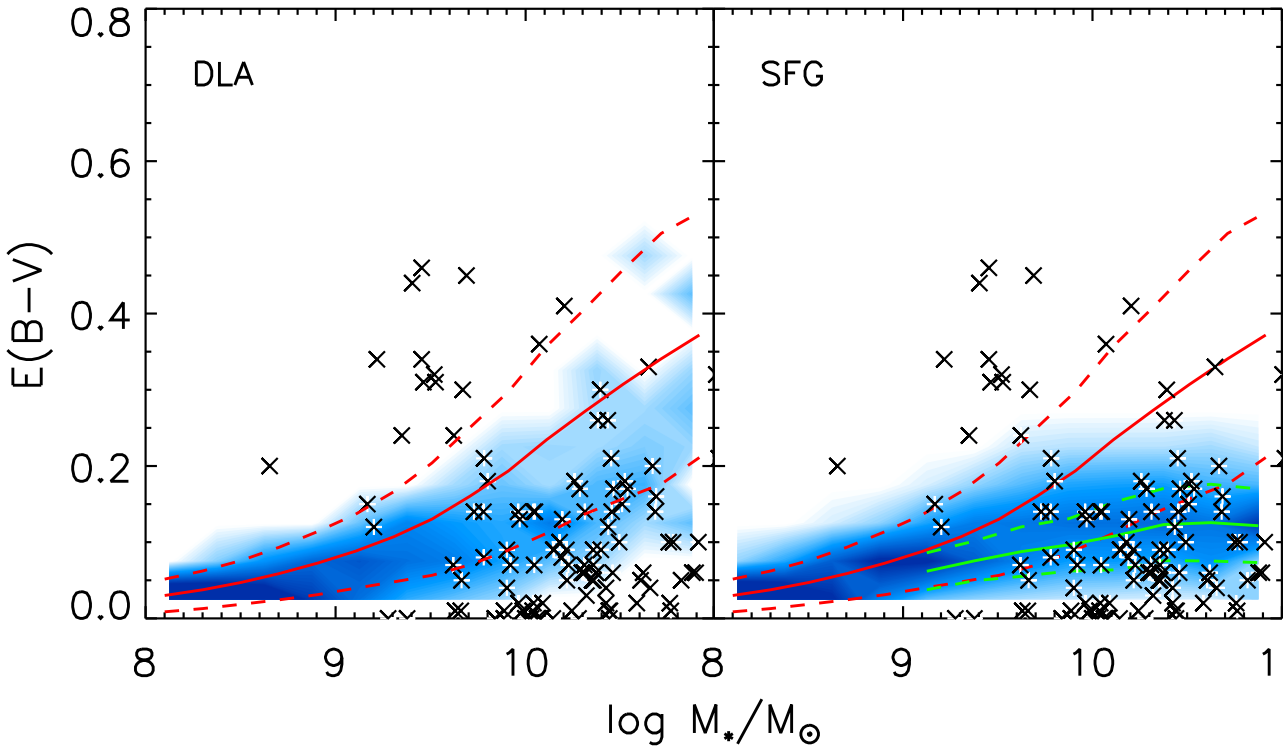

Figure 14. Same as Fig. 12 except the conditional probability distribution of $E(B-V)$ value for a given stellar mass is shown. Observational estimates of $E(B-V)$ for SFGs at $1.7 < z < 2.3$ from Reddy et al. (2006) are overplotted (crosses). At high stellar masses, DLA host galaxies and SFGs have lower $E(B-V)$ values at a given stellar mass than a typical galaxy in a stellar mass selected sample. Additionally, our magnitude-limited sample of SFGs with $R < 25.5$ preferentially selects galaxies with even lower $E(B-V)$ values at the same stellar mass.

surveys. In our models, galaxies selected as DLAS have a lognormal distribution in H I mass centred on $\log M_{\rm s}/{\rm M}_{\odot} \sim 9.5$. At lower masses, galaxies have an insufficient amount of cold gas to have a substantial DLA cross-section, and are increasingly unlikely to be selected as DLAS. It is important to note that we assume that cold gas is distributed as an exponential disc in all galaxies. It is possible that low-mass galaxies have alternative cold gas distributions that might increase their DLA cross-sections, making it possible that a higher fraction would be selected as DLAS.

Additionally, the majority of model DLAS exhibit similar colours to typical galaxies at the same magnitudes, implying they are not photometrically distinct. As most DLAS are quite faint, they generally exhibit blue colours. This indicates that the majority would be selected using common colour–colour selection techniques (*U*-band dropouts, BM/BXs, and/or BzKs) except they are too faint owing to their low stellar masses and SFRs. Only a small fraction are bright enough to be detected in traditional magnitude-limited surveys (e.g. $B_{435} < 26.5$, $R \leq 25.5$ or $K \leq 23$).

Our pressure-based BRj25 model predicts that DLA host galaxies are $\gtrsim$1 mag fainter in the *R* band and $\gtrsim$2 mag fainter in the *H* band than the metallicity-based GKj25 model. Due to these differences, our GKj25 (BRj25) model predicts that ∼52 per cent (∼24 per cent) of DLA host galaxies should be observable at depths of $H \leq 26$ (see Fig. 6). These estimates do not take into account the proximity of the DLA host galaxy to the background quasar which will undoubtedly lower this fraction, as both models predict that ∼60 per cent are within 1 arcsec of the line of sight. However, more massive, more luminous DLAS are more likely to have larger impact parameters. The faintest have *H*-band magnitudes of $H \gtrsim 30$ and will require the next generation of optical/NIR telescopes such as *JWST*. By restricting our DLA sample to those with $\log Z > -1$, our model predicts that the rate of finding DLA counterparts should increase by ∼50 per cent, although this will preferentially identify DLAS with large stellar masses. This follows from the mass–metallicity relationship that generally holds for DLA hosts combined with the tight correlation between stellar mass and *H*-band magnitude. It is also in agreement with a correlation between DLA velocity width and stellar mass that has been seen by several studies (e.g. Rafelski et al. 2012; Neeleman et al. 2013), and shown to hold in our models in B14.

Our models predict that a large fraction of DLAS originate in low-mass galaxies, and as a result it is unsurprising that the majority suffer from only small amounts of dust reddening ($E(B-V) < 0.1$). With a magnitude limit of $H < 26$ on the host galaxy, a significant number of DLAS have higher $E(B-V)$ values; ∼43 per cent have $E(B-V) > 0.1$, owing to a substantial number originating in higher mass, more luminous galaxies. At these magnitudes, the distribution of DLA $E(B-V)$ values closely matches that of galaxies identified in a similar magnitude-limited survey of $H < 26$, indicating that bright DLA host galaxies at a given magnitude do not contain unusually low amounts of dust. The $E(B-V)$ values predicted by our models represent the average across the entire galaxy and assume a mixed dust ('slab') geometry. Our predictions are not directly comparable to results from stacking quasar lines of sight or radio-selected quasars, as both measure the amount of extinction along the DLA line of sight which is better represented by a screen geometry. None the less, the amount of dust reddening in DLA host galaxies in our models is generally consistent with the low $E(B-V)$ values observed in DLAS.

6.2 Line-of-sight properties versus galaxy-wide averages

In addition to the broad range of properties exhibited by DLA host galaxies, our models suggest that there is a significant difference between the properties along the DLA line of sight as compared to the mass-weighted average over the host galaxy. Recently, observations of a few DLA host galaxies (Krogager et al. 2012, 2013) have allowed us to probe this connection. One insight arising from this work is the success of our simple model in reproducing the observed impact parameter distribution of DLAS as well as the relation between impact parameter, column density, and metallicity. Our models successfully predict that the majority of DLA host galaxies should be found within ∼1–2 arcsec of the quasar, indicating that the majority of DLAS at $z \sim 2$ are closely associated with galactic discs. Moreover, the metallicity-based GKj25 model reproduced the observed distribution of impact parameters as a function of column density and metallicity quite well. Observations and both of our models find a decreasing number of high column density DLAS at larger impact parameters, and a decreasing number of DLAS at a given impact parameter with increasing column density. These two effects are consistent with DLAS arising in discs, with high column density systems originating in the central regions of a galaxy where cold gas densities are largest. Additionally, as the cold gas density decreases with radius, DLAS at large impact parameters must have low column densities. Note that galaxies with clumpy interstellar media or large infalling clumps could potentially produce high-$N_{\rm H\,I}$ DLAS at large impact parameters. In spite of these relations, our model predicts the majority of DLAS will be found at low column density and low impact parameter due to the large contribution of low-mass DLA host galaxies with small DLA cross-sections. These types of objects are not present in the observational sample, but this may be due to selection effects.

Additionally, our models find that only the highest metallicity systems are found at large impact parameters, indicating that only the most massive galaxies have sufficient cold gas reservoirs to have large enough DLA cross-sections to host DLAS with $b > 20$ kpc. However, when we implement a modest metallicity gradient of $\mathrm{d}Z/\mathrm{d}r = -0.1$ dex kpc^{-1}, both our models predict that DLAS at high impact parameter will be more metal poor. These high impact parameter, low-metallicity galaxies are not seen in observations and may indicate that metallicity gradients do not significantly contribute to the low metallicities observed in DLAS.

Our models find it difficult to explain the existence of high column density, low-metallicity DLAS. Our exponential cold gas discs have a decreasing gas density with radius, so in our models, these systems must arise in the central regions of galaxies. However, the highest column density systems tend to arise in more massive galaxies where there is more cold gas and more metals. These systems may be impacted by cold 'pristine' gas, such as in clumps or streams, increasing the column density while also diluting the metallicity. Our models do predict the presence of high metallicity, high column density systems, which have not been observed. The absence of these systems may be a selection effect as the high metallicities and column densities may result in more dust reddening, causing them to drop out of colour-selected quasar surveys. The presence of a metallicity gradient would increase the number of these systems as metallicities are typically higher in the central regions of galaxies where column densities are also higher.

Krogager et al. (2012) found an increased number of host galaxy counterparts to high-metallicity DLAS, supporting the idea that DLAS follow a mass–metallicity relation. Our models predict that selecting high metallicity DLAS will also bias the impact parameter distribution to larger values for this same reason. Yet, the presence of a metallicity gradient would decrease the average impact parameter of high-metallicity DLAS as they would sample the central regions of the galaxy. These effects are clearly non-trivial, speculative, and likely vary greatly from galaxy to galaxy; therefore a detailed examination of DLAS at large impact parameters including the distribution and composition of cold gas at those radii would be necessary to decompose these effects. While only a small number of DLA host galaxies have been identified, it is intriguing that we find that DLAS arising from exponential discs are generally consistent with the trends observed between DLA column density, metallicity, and size.

As DLAS preferentially probe gas in the outer parts of galaxies, the properties of DLAS as measured along the line of sight likely differ substantially from the central regions of the host galaxy. Using the cold gas density at the location of the DLA sightline, we calculate the H_2 fraction and the SFR as would be measured for the DLA. Our models predict that in most cases, the H_2 fraction and SFR along the DLA line of sight are substantially lower than in the central regions. The small scale radii of the H_2 discs causes the majority of the mass of H_2 to be in the central regions, leading to low predicted H_2 fractions in DLAS. The lower H_2 fractions along the DLA line of sight yield substantially lower SFRs of ∼2 decades at the DLA sightline than would be measured in the central regions of the host galaxy (see e.g. Fig. 9).

Due to the large range of DLA host galaxy masses and impact parameters, the SFR distribution spans several decades. The width of this distribution is comparable for the DLA sightlines, offering a possible explanation for why some DLAS show no C II* absorption, indicating SFRs of log SFR $\lesssim -3$, while others exhibit significant C II*. In rare cases, where the DLA line of sight passes through the central regions of the galaxy, the H_2 fraction and SFR measured in the DLA may be larger than the average for the galaxy. These rare circumstances would provide an opportunity to study the relations between cold gas density, H_2 fraction, and SF in high-redshift galaxies.

Finally, the metallicity probed by emission lines has been observed to be a few tenths of a dex higher than that measured in the DLA (Bouché et al. 2013). Our implementation of a modest metallicity gradient of $\mathrm{d}Z/\mathrm{d}r = 0.1$ dex kpc^{-1} yields an average shift in metallicity of $\Delta Z = -0.26$, consistent with observations. Our model predicts average metallicities in SFGs that are consistent with observations at $z \sim 2$. Moreover, implementing a modest metallicity gradient reproduces the metallicities observed in DLAS at the same redshift. As observations of SFGs at $z \sim 2$ display a large range of metallicity gradients, it is unlikely that all galaxies follow this simplified prescription. However on average, the presence of a metallicity gradient may contribute to the low metallicities observed in DLAS.

6.3 DLAS, SFGs, and galaxy scaling relations

At $z \sim 2$, we find both DLAS and SFGs to generally follow the trends between DLA cross-section, cold gas fraction, SFR, metallicity, and stellar mass that are exhibited by the underlying galaxy population. At a given stellar mass, DLA host galaxies have higher DLA cross-sections compared to the general galaxy population at the same stellar mass, with a larger offset at lower stellar masses (i.e. a flatter $\sigma_{\rm DLA} - M_{\rm s}$ relation). As DLAS are selected based on the DLA cross-section and number density of galaxies, the decreasing DLA cross-sections with stellar mass and increasing number of low-mass systems roughly balance out.

Selecting SFGs with bright rest-frame ultraviolet continua ($R < 25.5$) identifies galaxies with larger DLA cross-sections, where again this difference is larger at the low stellar mass end. This magnitude limit effectively selects SFGs with a minimum SFR and therefore a minimum cold gas mass. At lower stellar masses, these galaxies have the same amount of cold gas, making their DLA cross-sections larger than the typical SFG that is not subject to this magnitude limit. Therefore, these galaxies are also more likely to be associated with DLAS. When studying the relation between cold gas fraction and stellar mass, we find similar trends for both DLAS and SFGs. Due to the large cold gas masses associated with DLA host galaxies, at low stellar masses they are almost entirely composed of cold gas. The SF threshold imposed by a rest-UV magnitude limit causes low-mass SFGs with $R < 25.5$ to have extremely high cold gas fractions. However, SFGs with a magnitude limit of $R < 25.5$ have stellar masses $\gtrsim 10^9$ M$_\odot$, while DLAS can arise from galaxies with masses down to $\sim 10^6$ M$_\odot$.

DLA host and SFGs generally follow the SFR–stellar mass relation of the entire galaxy population, where both are fairly accurate tracers of the SFR at a given stellar mass. DLA host galaxies do tend to have higher SFRs at low stellar mass relative to normal galaxies. The magnitude limit of $R < 25.5$ identifies galaxies with a minimum SFR that causes lower mass galaxies to have higher SFRs relative to their stellar mass, although at higher stellar masses this approaches the SFRs of 'typical' (stellar mass selected) galaxies. Observations of SFGs at the same redshift indicate that our model galaxies have SFRs that are too low at a given stellar mass. This is a common problem in cosmological models of galaxy formation and may be linked with the difficulty in correctly reproducing the SF histories of galaxies in such models (e.g. Somerville & Davé 2015; White et al. 2015). This discrepancy remains an open challenge for modellers (though see White et al. 2015 and Henriques et al. 2013 for proposed solutions).

Another insight from our study is that both DLA host and SFGs follow the mass–metallicity relation, which is consistent with observations. However, if DLA metallicity is correlated with the distance from the host galaxy as in the case of a metallicity gradient, then our models predict that high-mass DLAS will exhibit a broad range in metallicities. A modest metallicity gradient of $\mathrm{d}Z/\mathrm{d}r = -0.1$ dex kpc^{-1} will effectively wipe out any trend between stellar mass and metallicity above $\log M_\mathrm{s}/\mathrm{M}_\odot \gtrsim 9$ due to the large range in impact parameters for massive galaxies. Below this mass, DLAS generally follow the mass–metallicity relation, albeit with a larger scatter. Therefore depending on how prevalent metallicity gradients are, inferring the mass of a DLA based on its metallicity should be treated with caution. We find that both DLAS and SFGs are offset to slightly lower metallicities at a given stellar mass than the overall population, which is consistent with the Fundamental Mass–Metallicity Relation, in which galaxies with higher SFR (and higher gas fractions) are observed to have systematically lower metallicities at a given stellar mass (Mannucci et al. 2010; Bothwell et al. 2013; Lara-López et al. 2013).

We find that similar to typical galaxies, at low stellar masses DLA host galaxies have not undergone enough SF to have created a substantial amount of metals, resulting in small amounts of dust extinction. The amount of dust extinction increases with stellar mass, causing high mass DLA host galaxies to be more reddened and therefore likely to be missed from colour-selected quasar surveys. The lack of observed high column density, high metallicity DLAS further supports this hypothesis. None the less, DLA host galaxies exhibit less dust reddening than typical galaxies at the same stellar mass. This difference is caused by their lower metallicities. If DLAS probe the outer regions of galaxies where the metallicities are lower, they should have even lower amounts of dust extinction. Unsurprisingly, we find that model SFGs selected to be luminous in the rest-frame ultraviolet do not contain significant amounts of dust reddening at any stellar mass.

6.4 Theoretical interpretation and implications for models

The nature of DLAS and their relationship with galaxies still pose an open puzzle. On the one hand, DLAS have a similar clustering strength to that of fairly luminous SFGs (Bouché et al. 2005; Font-Ribera et al. 2012), suggesting that they should be hosted by haloes with similar masses. Yet, despite extensive searches, many DLAS have not yielded detections of optical counterparts (Warren et al. 2001; Krogager et al. 2012), suggesting that at least some DLAS are not closely associated with luminous SFGs.

Our standard disc models (in which the specific angular momentum of the baryons in the disc is the same as that of the dark matter halo) have been extensively tested and calibrated against observations, and shown to reproduce many of the properties of galaxies as traced by their stellar content and interstellar gas detected in emission (PST14, SPT15). However, we showed in B14 that the otherwise very successful standard disc model did not fare at all well at reproducing the observed properties of DLAS (under the assumption that DLAS arise from the ISM within galactic discs). Therefore we introduced the 'extended disc' model, in which we assumed that the gas that forms a disc possesses 2.5 times more specific angular momentum than the dark matter halo. This assumption was motivated by results from recent hydrodynamic simulations, and resulted in much greater success at reproducing the observed column density distribution of DLAS at $z \lesssim 2.5$, DLA host halo masses (as constrained by clustering), DLA kinematics, and DLA metallicities. In this paper, we have shown that the extended disc model is also very successful at reproducing the impact parameter distribution for DLA hosts and correlations between impact parameter and column density and metallicity. Our extended disc model also provides a neat explanation of the paradox above: because DLAS are hosted by haloes and galaxies with a very broad range of masses and luminosities, they can cluster like SFGs but, a significant fraction of the time, no SFG will be detected nearby, at the depth of existing observational studies.

Yet the alert reader may be concerned about whether the extended disc model still reproduces the observational constraints for galaxies detected in emission. We showed a comparison of the main calibration quantities at $z = 0$ for the standard and extended disc models in fig. 1 of B14. To summarize the results shown there: the extended disc model produces a sharper cutoff at the high-mass end of the stellar mass function (hence fewer massive galaxies), but below the 'knee' in the mass function the results are nearly identical. The flip side of this is that there are more galaxies with large H I masses in the extended disc model. The H2MF is nearly the same in both models. We have verified that similar results hold at $z = 2$: the extended disc model produces somewhat fewer galaxies with large stellar masses ($m_\mathrm{star} > m_\mathrm{char}$, where m_char is the characteristic mass of the 'knee' of the SMF), but the two models produce near identical predictions below m_char. The HIMFs and H2MFs are extremely similar for the two models. The ratio of H I and H_2 mass to stellar mass is higher in the extended disc model, but the slope of the dependence of both quantities on stellar mass remains the same. Overall, the differences between the two models are minor enough that the extended disc models can be retuned to give an equally good match to the calibration quantities (although we have

not done this here). At $z = 2$, the focus of this paper, the stellar mass function predicted by the extended disc model is within the error bars on the available observations. Similarly, we have checked that the predictions shown in Figs 1, 2, 3, and 4 for the extended disc model in this paper, are sufficiently similar for the standard disc model, that none of our conclusions would change were we to adopt it.

The one exception is the disc mass–size relationship. We have shown previously that our standard disc model reproduces the observed mass–size relation from $z \sim 0$–2 for stellar discs and for H_2 discs (as constrained by CO observations or H α observations; see PST14). However, our extended disc model produces stellar and H_2 discs that are about a factor of 2 too large compared with observations at a given stellar mass, at all redshifts $z \sim 0$–2. Moreover, we find that the size–mass relation predicted by our standard disc model agrees with that predicted by the high-resolution numerical hydrodynamic simulations of Christensen et al. (2012), while the extended disc model produces larger discs (Berry 2014).

We have noted before that several studies based on numerical hydrodynamic simulations have shown that the gas that forms the baryonic disc has a mean specific angular momentum that is about a factor of 2–3 larger than that of the dark matter halo (Brooks et al. 2009; Stewart et al. 2013; Danovich et al. 2015). However, the recent study of Danovich et al. (2015) also showed that after being accreted into the inner halo, the gas loses about half of its angular momentum due to torques induced by misalignment with the inner disc, and by violent disc instabilities. Danovich et al. (2015) note that after a few orbital times the resulting baryonic disc has a specific angular momentum that is similar to that of the dark matter, in keeping with the simple assumption of the standard disc model, but this good agreement is actually just a coincidence. It is also intriguing that Danovich et al. (2015) find extended rings of cold gas in their simulations, with spin parameters roughly twice those of the halo, and with H I column densities high enough to result in DLAS for 30 per cent of all lines of sight. These are remarkably reminiscent of our extended gas discs. What is clearly not being treated properly in our current model is the loss of angular momentum by the baryons that will end up making the stellar and H_2 disc during the final phases of the infall process. It is still unclear whether the relationship between the angular momentum of the halo and that of the baryonic disc will be a function of galaxy mass, cosmic time, merger history, or other parameters, and how sensitive these details will be to the specific implementation of sub-resolution physics in the numerical simulations. We are currently working on a detailed analysis of high-resolution numerical hydrodynamic simulations from several different codes and groups in order to better understand these issues.

Readers may also be concerned that if the sizes of H_2 discs in our models are incorrect, then all predicted properties of galaxies will be off because the density of H_2 gas determines the SF efficiency, through the KS relation. It turns out that this is less true than one might expect, for the reason that SF is highly self-regulated in our models. More dense gas leads to a higher SFR, which leads to stronger feedback and less dense gas in the next timestep (due to outflows and consumption), which reduces the SFR in the next timestep (see the extensive discussion in Somerville & Davé 2015). Therefore, as shown explicitly in White et al. (2015) and SPT15, a changing SF efficiency has a weak effect on the stellar mass function and most other predicted quantities. This is the reason that we saw so little change in the global quantities predicted in our standard and extended disc models, even though the gas densities and hence the SF efficiencies were quite different in those two models.

7 CONCLUSIONS

Using SAMs of galaxy formation, we study the properties of DLA host galaxies and their relationship to SFGs at $z \sim 2$. Our preferred model, GKj25, uses a metallicity-based prescription for H_2 formation and an 'extended' gas distribution as parametrized by f_j, the ratio of cold gas specific angular momentum to dark matter specific angular momentum. This work builds upon the results of B14, who showed that this model was the most successful at reproducing the column density distribution, comoving number density, metallicity distribution, and velocity width distribution of DLAS at $z \lesssim 2.5$. We began by comparing a sample of model SFGs to observed ones by 'observing' model galaxies that were identified using the same selection criteria. After verifying that model SFGs have predicted statistical properties similar to observed ones, we characterized the properties of DLA host galaxies, studied the relationship between DLAS and their hosts, and compared the properties of DLA host galaxies to SFGs.

Our main results are as follows.

(i) DLA host galaxies exhibit a broad range of galaxy properties extending over several decades in stellar mass, SFR, cold gas mass, and luminosity. As a result, they are equally likely to arise from low- and high-mass galaxies (in terms of total or stellar mass). This also implies that high-mass galaxies are much more likely to host DLAS due to the smaller number of high-mass galaxies relative to low mass ones. Luminous, high-mass DLA host galaxies also have the largest amounts of dust extinction, which may bias colour-selected QSO surveys.

(ii) The photometric colours of DLA host galaxies are not significantly biased relative to objects with similar stellar masses at the same redshift. However, the majority of DLAS are much fainter than common magnitude limits for optical surveys. Obtaining complete samples will likely require the next generation of IR sensitive telescopes such as *JWST*. However, in the meantime, it may be feasible to significantly enlarge our samples of DLA hosts with deep imaging with WFC3 on *HST*.

(iii) The predicted distribution of impact parameters of model DLAS are consistent with observations of DLAS with identified host galaxy counterparts. The relation between impact parameter and column density and metallicity is consistent with DLAS arising from cold gas discs, however, metallicity gradients predict more high impact parameter, low metallicity systems than are observed. The H_2 fraction, SFR, and metallicity as measured along the DLA line of sight likely differ substantially from the host galaxy, which may help explain the low SFRs and metallicities observed in DLAS.

(iv) In general, DLA host galaxies and SFGs follow the same relations between stellar mass and DLA cross-section, cold gas fraction, SFR, metallicity, and dust extinction as observed in the overall galaxy population. DLAS preferentially select galaxies with more cold gas, which causes them to have higher cold gas fractions, higher SFRs, lower metallicities, and less dust extinction at a given stellar mass than the overall galaxy population.

ACKNOWLEDGEMENTS

We thank Xavier Prochaska for stimulating discussions and Marc Rafelski for comments on the manuscript. RSS thanks the Downsbrough family for their generous support. EG acknowledges support from NSF Career Grant 1055919. AM acknowledges support from NSF Grant 1153335.

REFERENCES

Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379
Berry M. J., 2014, PhD thesis, Univ. New Jersey
Berry M., Somerville R. S., Haas M. R., Gawiser E., Maller A., Popping G., Trager S. C., 2014, MNRAS, 441, 939 (B14)
Bigiel F., Blitz L., 2012, ApJ, 756, 183
Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846
Blitz L., Rosolowsky E., 2004, ApJ, 612, L29 (BR)
Blitz L., Rosolowsky E., 2006, ApJ, 650, 933
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27
Bondi H., 1952, MNRAS, 112, 195
Bothwell M. S., Maiolino R., Kennicutt R., Cresci G., Mannucci F., Marconi A., Cicone C., 2013, MNRAS, 433, 1425
Bouché N., Gardner J. P., Katz N., Weinberg D. H., Davé R., Lowenthal J. D., 2005, ApJ, 628, 89
Bouché N., Murphy M. T., Kacprzak G. G., Péroux C., Contini T., Martin C. L., Dessauges-Zavadsky M., 2013, Science, 341, 50
Boylan-Kolchin M., Ma C.-P., Quataert E., 2008, MNRAS, 383, 93
Bromm V., Larson R. B., 2004, ARA&A, 42, 79
Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, ApJ, 694, 396
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cen R., 2012, ApJ, 748, 121
Chabrier G., 2003, PASP, 115, 763
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Christensen C., Quinn T., Governato F., Stilp A., Shen S., Wadsley J., 2012, MNRAS, 425, 3058
Cooke J., Wolfe A. M., Gawiser E., Prochaska J. X., 2006, ApJ, 652, 994
Daddi E., Cimatti A., Renzini A., Fontana A., Mignoli M., Pozzetti L., Tozzi P., Zamorani G., 2004, ApJ, 617, 746
Dahlen T. et al., 2013, ApJ, 775, 93
Danovich M., Dekel A., Hahn O., Ceverino D., Primack J., 2015, MNRAS, 449, 2087
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Ellison S. L., Yan L., Hook I. M., Pettini M., Wall J. V., Shaver P., 2001, A&A, 379, 393
Ellison S. L., Hall P. B., Lira P., 2005, AJ, 130, 1345
Elmegreen B. G., 1989, ApJ, 338, 178
Elmegreen B. G., 1993, ApJ, 411, 170
Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006, ApJ, 644, 813
Flores R., Primack J. R., Blumenthal G. R., Faber S. M., 1993, ApJ, 412, 443
Font-Ribera A. et al., 2012, J. Cosmol. Astropart. Phys., 11, 59
Fontana A. et al., 2006, A&A, 459, 745
Fontanot F., De Lucia G., Monaco P., Somerville R. S., Santini P., 2009, MNRAS, 397, 1776
Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J. R., 2011, MNRAS, 418, 1796
Fumagalli M., O'Meara J. M., Prochaska J. X., Rafelski M., Kanekar N., 2015, MNRAS, 446, 3178
Fynbo J. P. U. et al., 2010, MNRAS, 408, 2128
Fynbo J. P. U. et al., 2011, MNRAS, 413, 2481
Glover S., 2013, in Wiklind T., Mobasher B., Bromm V., eds, Astrophysics and Space Science Library, Vol. 396, The First Galaxies. Springer-Verlag, Berlin, p. 103
Gnedin N. Y., 2000, ApJ, 542, 535
Gnedin N. Y., 2012, ApJ, 754, 113
Gnedin N. Y., Kravtsov A. V., 2010, ApJ, 714, 287
Gnedin N. Y., Kravtsov A. V., 2011, ApJ, 728, 88 (GK)
Grazian A. et al., 2007, A&A, 465, 393
Greif T. H., Glover S. C. O., Bromm V., Klessen R. S., 2010, ApJ, 716, 510
Grogin N. A. et al., 2011, ApJS, 197, 35
Haiman Z., Rees M. J., Loeb A., 1996, ApJ, 467, 522
Hathi N. P. et al., 2010, ApJ, 720, 1708
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, MNRAS, 431, 3373
Hopkins P. F. et al., 2009a, MNRAS, 397, 802
Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009b, ApJ, 691, 1168
Jorgenson R. A., Wolfe A. M., Prochaska J. X., Lu L., Howk J. C., Cooke J., Gawiser E., Gelino D. M., 2006, ApJ, 646, 730
Kajisawa M. et al., 2009, ApJ, 702, 1393
Karim A. et al., 2011, ApJ, 730, 61
Kennicutt R. C., Jr, 1998, ApJ, 498, 541
Khare P., vanden Berk D., York D. G., Lundgren B., Kulkarni V. P., 2012, MNRAS, 419, 1028
Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102
Komatsu E. et al., 2011, ApJS, 192, 18
Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73
Kravtsov A. V., Gnedin O. Y., Klypin A. A., 2004, ApJ, 609, 482
Krogager J.-K., Fynbo J. P. U., Møller P., Ledoux C., Noterdaeme P., Christensen L., Milvang-Jensen B., Sparre M., 2012, MNRAS, 424, L1
Krogager J.-K. et al., 2013, MNRAS, 433, 3091
Krumholz M. R., McKee C. F., Tumlinson J., 2009, ApJ, 699, 850
Kulas K. R. et al., 2013, ApJ, 774, 130
Lara-López M. A. et al., 2013, MNRAS, 433, 35
Lu Y. et al., 2014, ApJ, 795, 123
Macciò A. V., Kang X., Fontanot F., Somerville R. S., Koposov S., Monaco P., 2010, MNRAS, 402, 1995
Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388
Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2011, A&A, 528, A35
Maiolino R. et al., 2008, A&A, 488, 463
Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, MNRAS, 408, 2115
Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, ApJ, 701, 1765
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Murphy M. T., Bernet M. L., 2016, MNRAS, 455, 1043
Murphy M. T., Liske J., 2004, MNRAS, 354, L31
Nakamura F., Umemura M., 2001, ApJ, 548, 19
Narayanan D., Krumholz M. R., Ostriker E. C., Hernquist L., 2012, MNRAS, 421, 3127
Neeleman M., Wolfe A. M., Prochaska J. X., Rafelski M., 2013, ApJ, 769, 54
Niemi S.-M., Somerville R. S., Ferguson H. C., Huang K.-H., Lotz J., Koekemoer A. M., 2012, MNRAS, 421, 1539
Noterdaeme P., Petitjean P., Ledoux C., Srianand R., 2009, A&A, 505, 1087
Noterdaeme P. et al., 2012, A&A, 547, L1
Oesch P. A. et al., 2010, ApJ, 725, L150
Ostriker E. C., Shetty R., 2011, ApJ, 731, 41
Pei Y. C., Fall S. M., Bechtold J., 1991, ApJ, 378, 6
Péroux C., McMahon R. G., Storrie-Lombardi L. J., Irwin M. J., 2003, MNRAS, 346, 1103
Pontzen A. et al., 2008, MNRAS, 390, 1349
Popping G., Somerville R. S., Trager S. C., 2014, MNRAS, 442, 2398 (PST14)
Porter L. A., Somerville R. S., Primack J. R., Johansson P. H., 2014, MNRAS, 444, 942
Prochaska J. X., Wolfe A. M., 2009, ApJ, 696, 1543
Qian Y.-Z., Wasserburg G. J., 2003, ApJ, 596, L9
Rafelski M., Wolfe A. M., Chen H.-W., 2011, ApJ, 736, 48
Rafelski M., Wolfe A. M., Prochaska J. X., Neeleman M., Mendez A. J., 2012, ApJ, 755, 89
Rahmati A., Schaye J., 2014, MNRAS, 438, 529
Reddy N. A., Erb D. K., Steidel C. C., Shapley A. E., Adelberger K. L., Pettini M., 2005, ApJ, 633, 748

Reddy N. A., Steidel C. C., Erb D. K., Shapley A. E., Pettini M., 2006, ApJ, 653, 1004
Reddy N. A., Steidel C. C., Pettini M., Adelberger K. L., Shapley A. E., Erb D. K., Dickinson M., 2008, ApJS, 175, 48
Robertson B., Cox T. J., Hernquist L., Franx M., Hopkins P. F., Martini P., Springel V., 2006a, ApJ, 641, 21
Robertson B., Bullock J. S., Cox T. J., Di Matteo T., Hernquist L., Springel V., Yoshida N., 2006b, ApJ, 645, 986
Schneider R., Ferrara A., Natarajan P., Omukai K., 2002, ApJ, 571, 30
Shapley A. E., Steidel C. C., Adelberger K. L., Dickinson M., Giavalisco M., Pettini M., 2001, ApJ, 562, 95
Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65
Shapley A. E., Steidel C. C., Erb D. K., Reddy N. A., Adelberger K. L., Pettini M., Barmby P., Huang J., 2005, ApJ, 626, 698
Sobral D., Best P. N., Matsuda Y., Smail I., Geach J. E., Cirasuolo M., 2012, MNRAS, 420, 1926
Somerville R. S., Davé R., 2015, ARA&A, 53, 51
Somerville R. S., Kolatt T. S., 1999, MNRAS, 305, 1
Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087
Somerville R. S., Primack J. R., Faber S. M., 2001, MNRAS, 320, 504
Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008a, MNRAS, 391, 481 (S08)
Somerville R. S. et al., 2008b, ApJ, 672, 776
Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, MNRAS, 423, 1992 (S12)
Somerville R. S., Popping G., Trager S. C., 2015, MNRAS, 453, 4337 (SPT15)
Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, ApJ, 462, L17
Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2003, ApJ, 592, 728
Steidel C. C., Shapley A. E., Pettini M., Adelberger K. L., Erb D. K., Reddy N. A., Hunt M. P., 2004, ApJ, 604, 534
Stewart K. R., Brooks A. M., Bullock J. S., Maller A. H., Diemand J., Wadsley J., Moustakas L. A., 2013, ApJ, 769, 74
Storrie-Lombardi L. J., Wolfe A. M., 2000, ApJ, 543, 552
Tacconi L. J. et al., 2010, Nature, 463, 781
Tacconi L. J. et al., 2013, ApJ, 768, 74
Vladilo G., Prochaska J. X., Wolfe A. M., 2008, A&A, 478, 701
Warren S. J., Møller P., Fall S. M., Jakobsen P., 2001, MNRAS, 326, 759
Whitaker K. E. et al., 2011, ApJ, 735, 86
White C. E., Somerville R. S., Ferguson H. C., 2015, ApJ, 799, 201
Wise J. H., Turk M. J., Norman M. L., Abel T., 2012, ApJ, 745, 50
Wolfe A. M., Chen H.-W., 2006, ApJ, 652, 981
Wolfe A. M., Prochaska J. X., 2000, ApJ, 545, 591
Wolfe A. M., Gawiser E., Prochaska J. X., 2005, ARA&A, 43, 861
Wolfe A. M., Prochaska J. X., Jorgenson R. A., Rafelski M., 2008, ApJ, 681, 881 (W08)

This paper has been typeset from a TEX/LATEX file prepared by the author.